

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 11:46:34 ; Search time 28.2264 Seconds
(without alignments)
1605.064 Million cell updates/sec

Title: US-10-021-121-4

Perfect score: 1850
Sequence: 1 MGPPIGSGPGRVRCALLLIG.....PYIVQDGPSPGPPIYKVV 340

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database :
1: A_Geneseq_101002:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
24: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1850	100.0	340	18	AAW33699
2	1850	100.0	340	18	AAW31544
3	1850	100.0	340	18	AAW10637
4	1844	99.7	340	18	AAW17081
5	1841	99.5	455	18	AAW33698
6	1836	99.2	340	19	AAW46615
7	637.5	34.5	334	17	AAW00287
8	637.5	34.5	336	17	AAW92742
9	632	34.2	346	15	AAW55059
10	632	34.2	346	17	AAW91930

11	632	34.2	346	18	AAW36055
12	632	34.2	346	18	AAW19249
13	632	34.2	346	18	AAW44233
14	631	34.1	346	16	AAW82606
15	630.5	34.0	331	17	AAW00288
16	629.5	34.0	333	17	AAW06337
17	629.5	34.0	333	17	AAW46555
18	629.5	34.0	333	17	AAW92743
19	629.5	34.0	333	17	AAW89287
20	629.5	34.0	333	17	AAW11308
21	619.5	33.5	308	17	AAW06334
22	619.5	33.5	308	17	AAW4656
23	492	26.6	89	21	AAW1438
24	453.5	24.5	658	21	AAW6782
25	453.5	24.5	683	21	AAW6781
26	447	24.2	229	23	AAW24019
27	443	23.9	229	23	AAW24020
28	431.5	23.3	195	17	AAW06333
29	431.5	23.3	195	18	AAW11307
30	284.5	15.4	92	22	AAW37671
31	284.5	15.4	92	23	ABG46524
32	205.5	11.1	136	22	AAW37534
33	205.5	11.1	136	23	ABG46394
34	204	11.0	82	21	AAW1437
35	202.5	10.9	106	21	AAW54187
36	201	10.9	82	21	AAW1436
37	182	9.8	652	22	ABW58621
38	182	9.8	652	22	ABW55548
39	181	9.8	234	16	AAW82605
40	179	9.7	238	16	AAW1481
41	176	9.5	184	17	AAW02587
42	176	9.5	184	17	AAW1283
43	176	9.5	184	17	AAW06820
44	176	9.5	209	17	AAW94766
45	176	9.5	209	19	AAW1006
46	175.5	9.5	213	20	AAW06822
47	172.5	9.3	218	21	AAW54058
48	172	9.3	200	17	AAW94767
49	172	9.3	200	19	AAW1007
50	169.5	9.2	201	16	AAW1482

ALIGNMENTS

RESULT 1
ID AAW33699 standard; Protein: 340 AA.

AC AAW33699;

DT 30-APR-1998 (first entry)

DE AL-2-short (AL-2s) protein.

XX AL-2s: AL-2; AL-2-short; human; treatment; neurological disorder; tumour;
XX rheumatoid arthritis; wound healing; paralysis; angiogenesis; leukaemia;
XX psoriasis; Alzheimer's disease; epilepsy.

OS Homo sapiens.

XX Key Location/Qualifiers
XX Peptide 1..26
XX Protein /note= "signal peptide"
XX Domain /note= "mature protein"
XX Domain /note= "extracellular domain"
XX Domain /note= "220..245"
XX Domain /note= "hydrophobic transmembrane domain"

XX W09740153-AL.

30-OCT-1997.
 17-APR-1997; 97WO-US06345.
 19-APR-1996; 96US-0635130.
 (GERTH) GENENTECH INC.
 Caras IW;
 WPI, 1997-535637/49.
 N-PSDB; AAW06355.
 Human AL-2 neurotrophic factor and related DNA - used to develop products for, e.g. treating neurologic disorders, angiogenesis disorders, tumours or rheumatoid arthritis or for wound healing

Claim 20; Fig 2A-B; 86pp; English.

This is a AL-2-short (AL-2s) protein. The AL-2 is a novel Eph-related tyrosine kinase receptor ligand. AL-2 can be administered to patients in whom the nervous system has been damaged by trauma, surgery, stroke, ischaemia, infection, metabolic disease, nutritional deficiency, malignancy, or toxic agents, to promote the survival or growth of neurons. They can be used to treat motoneuron disorders such as amyotrophic lateral sclerosis (Lou Gehrig's disease), Bell's palsy, and various conditions involving spinal muscular atrophy, or paralysis. AL-2 can be used to treat human neurodegenerative disorders, such as Alzheimer's disease, Parkinson's disease, epilepsy, demyelinating diseases such as multiple sclerosis, Huntington's chorea, Down's syndrome, nerve deafness, Menier's disease, and other disorders of the cerebellum. AL-2 can be used as cognitive enhancer, to enhance learning particularly in dementias or trauma, since they can promote axonal outgrowth and synaptic plasticity, particularly of hippocampal neurons that express AL-2 binding Eph-family receptors and cortical neurons that express AL-2. AL-2 can also be used for wound healing, i.e. accelerating neovascularisation of, e.g. burns and ulcers. The encoding nucleic acids are useful in preparing antibodies that specifically bind to the AL-2 protein. The antibodies and the AL-2 antagonists are useful in diagnosing and treating various neuronal disorders. AL-2 antagonists can be used for modulating angiogenesis. They can also be used for the treatment of tumours, acute myeloid leukaemia (AML), chronic myeloid leukaemia (CML), myelodysplastic syndrome (MDS), diabetic retinopathy, neovascular glaucoma, psoriasis and rheumatoid arthritis.

Sequence 340 AA:

Query Match 100.0%; Score 1850; DB 18; Length 340;
 Best Local Similarity 100.0%; Pred. No. 2.9e-143;
 Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MGPPHGPGVAVGALLLLGVLTGVLSLEPYWNSANKRFQAGGVLYPQIGRDL 60
 1 MGPPHGPGVAVGALLLLGVLTGVLSLEPYWNSANKRFQAGGVLYPQIGRDL 60
 61 LCPRAPPCGPHSSPNYFYKLYIVGAGGRCEAPPAVLLITCRPDIDLEFTIKFOY 120
 61 LCPRAPPCGPHSSPNYFYKLYIVGAGGRCEAPPAVLLITCRPDIDLEFTIKFOY 120
 121 SPMLWHEFRSHDYITINSDTREGLSLOGGVLTGMVYLRLVYGSPGGAVPKRP 180
 121 SPMLWHEFRSHDYITINSDTREGLSLOGGVLTGMVYLRLVYGSPGGAVPKRP 180
 121 SPMLWHEFRSHDYITINSDTREGLSLOGGVLTGMVYLRLVYGSPGGAVPKRP 180
 181 VSEMPERDRGAHSLPEKENVPGDPTSNATSRGAEGLPPSPMPAVAGAAGLALLLL 240
 181 VSEMPERDRGAHSLPEKENVPGDPTSNATSRGAEGLPPSPMPAVAGAAGLALLLL 240
 181 VSEMPERDRGAHSLPEKENVPGDPTSNATSRGAEGLPPSPMPAVAGAAGLALLLL 240
 241 GVAGAGAMWRRRAKPPSSSRPPGSGRGSGLGGCGMGPREAPPGGLGIALRG 300
 241 GVAGAGAMWRRRAKPPSSSRPPGSGRGSGLGGCGMGPREAPPGGLGIALRG 300
 301 GAADPPFCPHEKVGSDYGHPIYIVDGGPQSPSPNIYRY 340

Db 301 GAADPPFCPHEKVGSDYGHPIYIVDGGPQSPSPNIYRY 340

RESULT 2
 AAW31544
 ID AAW31544 standard; Protein; 340 AA.
 AAW31544;
 14-APR-1998 (first entry)
 Human cytokine Lerk-8.
 Lerk-8; cytokine; human; hek; elk; receptor tyrosine kinase; ligand; neurodegenerative disease; wound healing; neovascularisation; diagnosis; therapy.
 Homo sapiens.

Key Location/Qualifiers
 Peptide 1..27
 Protein /label= Sig_peptide 28..340
 Domain /label= Mat_protein 28..224
 Domain /note= "extracellular domain" 225..251
 Domain /note= "transmembrane domain" 252..340
 Modified-site /note= "cytoplasmic domain" 210..212
 MISC-difference 325 /note= "N-glycosylated"
 MISC-difference 325 /note= "residue 325 is Leu in Lerk-8 variant"

WO9736919-A2.
 09-OCT-1997.
 19-MAR-1997; 97WO-US04533.
 21-MAR-1996; 96US-0621146.
 (IMMV) IMMUNEX CORP.
 Cerrectl DP;
 WPI, 1997-503043/46.
 N-PSDB; AAT89519.

New isolated cytokine, Lerk-8 - binds to the hek and elk receptor tyrosine kinases, used to develop products for diagnosis and therapy
 Claim 3; Page 32-33; 37pp; English.

This protein sequence comprises a novel human cytokine designated Lerk-8. The amino acid sequence was deduced from a human foetal brain cDNA clone (see AAT89519). Lerk-8 binds to the cell surface receptors hek and elk, which are members of the eph/elk family of receptor tyrosine kinases. Lerk-8 polypeptides, especially soluble polypeptides comprising amino acid residues -27 to 142-197 of the full-length protein, can be expressed in transformed host cells. These polypeptides can be used to purify hek or elk proteins, and such cells can be used to identify cells that express hek or elk on the surface. Such cells can be used in various in vitro studies or in vivo procedures, e.g. neural cells expressing elk can be administered to a mammal afflicted with a neurodegenerative disorder. The Lerk-8 polypeptides can also be used to deliver diagnostic or therapeutic agents to these cells (e.g. leukaemia cells). The Lerk-8 DNA and polypeptides can also be used to treat disorders mediated by defective or insufficient amounts of Lerk-8; to treat disorders such as injury to neural tissue or neurologic disease, to promote angiogenesis, and for wound healing or stimulating

CC neovascularisation of grafted tissues.
XX
SQ Sequence 340 AA;

Query Match 100.0%; Score 1850; DB 18; Length 340;
Best Local Similarity 100.0%; Pred. No. 2.9e-143;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPPHSGGVRVGAALLLVGLVSGLSLEPYWNSANKRFOAEGGYVLYPQIGRLDL 60
DB 1 MGPPHSGGVRVGAALLLVGLVSGLSLEPYWNSANKRFOAEGGYVLYPQIGRLDL 60
QY 61 LCPRARPPGPHSSPNVEFYKLYLVGAQGRCEAPPAENLLTCDRDLRLFTIKFOEY 120
DB 61 LCPRARPPGPHSSPNVEFYKLYLVGAQGRCEAPPAENLLTCDRDLRLFTIKFOEY 120
QY 121 SPNLMGHEFRSHHDYIIATSDGTREGLESLOGVCCTTRGMKYLRLVQSPRGAVPRKP 180
DB 121 SPNLMGHEFRSHHDYIIATSDGTREGLESLOGVCCTTRGMKYLRLVQSPRGAVPRKP 180
QY 181 VSEWPMERDRGAASLEPGKENTPGDPTSNATSRGAEGPLPPSPMPAVAGAAGLALLL 240
DB 181 VSEWPMERDRGAASLEPGKENTPGDPTSNATSRGAEGPLPPSPMPAVAGAAGLALLL 240
QY 241 GVAAGAGAMCWRRRARAPSRHRPFGSGSLGLGGGGMGPREABPGEIGIALRGG 300
DB 241 GVAAGAGAMCWRRRARAPSRHRPFGSGSLGLGGGGMGPREABPGEIGIALRGG 300
QY 301 GAADPPPCPHYEKVSGDYGHPVYIVODGPPQSPNNITYKV 340
DB 301 GAADPPPCPHYEKVSGDYGHPVYIVODGPPQSPNNITYKV 340

RESULT 3
AAW10637
ID AAW10637 standard; Protein; 340 AA.

XX AAW10637;
XX 23-JUN-1997 (first entry)
XX
XX NLERK2 ligand for eph-related kinase.
XX
XX LERK: ligand for eph-related kinase; ERK; NLERK2;
XX receptor protein tyrosine kinase; cell proliferation;
XX cell differentiation; cell survival; nerve cell.
XX
XX Homo sapiens.
XX

FX Key Location/Qualifiers
FT Peptide 1..29
FT Protein /label= Sig_peptide
FT 30..340
FT Domain /label= Mat_protein
FT 227..251
FT Modified-site /label= Transmembrane_domain
FT 210
FT /label= N-glycosylation_site

PN WO9704091-A1.
XX
PD 06-FEB-1997.
XX
XX 19-JUL-1996; 96WO-AU00460.
XX
XX 05-FEB-1996; 96AU-0007890.
XX 20-JUL-1995; 95AU-0004263.
XX 27-NOV-1995; 95AU-0006847.
XX 22-DEC-1995; 95AU-0007299.
XX
XX (AMRA-) AMRAD OPERATIONS PTY LTD.
XX
XX Nicola NA;
PI

XX WPI, 1997-133632/12.
DR N-PSDB; AAT60966.
XX

PT Nucleic acid mol. encoding ligand for eph-related kinase - useful
PT for treatment of, pref. neuronal, cells to increase survival,
PT proliferation and differentiation
XX

PS Claim 16; Page 37-39; 71pp; English.

XX A novel human ligand for eph-related kinase (LERK) is designated
CC NLERK2 (AAW10637). It is encoded by a cDNA clone (AAT60966) obtd.
CC from a human foetal brain cDNA library. The novel receptor ligand
CC can be expressed in transformed host cells and used in methods
CC for regulating the development, maintenance or regeneration of
CC different cells (e.g. neurons) and tissues in vivo and in vitro.
CC Soluble NLERK2 peptides can be used to treat injury, disease or
CC abnormality in the nervous system, and membrane-bound NLERK2 to
CC modulate proliferation, different or survival e.g. in grafting
CC procedures or transplantation. NLERK2 can also be used to raise
CC antibodies for use in immunotherapy, and to detect anti-NLERK2
CC antibodies that may occur in some autoimmune diseases.

XX Sequence 340 AA;

Query Match 100.0%; Score 1850; DB 18; Length 340;
Best Local Similarity 100.0%; Pred. No. 2.9e-143;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPPHSGGVRVGAALLLVGLVSGLSLEPYWNSANKRFOAEGGYVLYPQIGRLDL 60
DB 1 MGPPHSGGVRVGAALLLVGLVSGLSLEPYWNSANKRFOAEGGYVLYPQIGRLDL 60
QY 61 LCPRARPPGPHSSPNVEFYKLYLVGAQGRCEAPPAENLLTCDRDLRLFTIKFOEY 120
DB 61 LCPRARPPGPHSSPNVEFYKLYLVGAQGRCEAPPAENLLTCDRDLRLFTIKFOEY 120
QY 121 SPNLMGHEFRSHHDYIIATSDGTREGLESLOGVCCTTRGMKYLRLVQSPRGAVPRKP 180
DB 121 SPNLMGHEFRSHHDYIIATSDGTREGLESLOGVCCTTRGMKYLRLVQSPRGAVPRKP 180
QY 181 VSEWPMERDRGAASLEPGKENTPGDPTSNATSRGAEGPLPPSPMPAVAGAAGLALLL 240
DB 181 VSEWPMERDRGAASLEPGKENTPGDPTSNATSRGAEGPLPPSPMPAVAGAAGLALLL 240
QY 241 GVAAGAGAMCWRRRARAPSRHRPFGSGSLGLGGGGMGPREABPGEIGIALRGG 300
DB 241 GVAAGAGAMCWRRRARAPSRHRPFGSGSLGLGGGGMGPREABPGEIGIALRGG 300
QY 301 GAADPPPCPHYEKVSGDYGHPVYIVODGPPQSPNNITYKV 340
DB 301 GAADPPPCPHYEKVSGDYGHPVYIVODGPPQSPNNITYKV 340

RESULT 4
AAW17081
ID AAW17081 standard; Protein; 340 AA.

XX AAW17081;
XX 09-AUG-1997 (first entry)
XX
XX Eph family ligand Efl-6.
XX
XX Efl-6; Eph; Etk; receptor tyrosine kinase; signal transduction;
XX ligand; neurological disease.
XX
XX Homo sapiens.
XX

FX Key Location/Qualifiers
FT Peptide 1..24
FT Protein /label= Sig_peptide
FT 25..340

```

FT FT /label= Mat_protein
FT Domain 225..249
FT /label= Transmembrane_domain
FT Misc-difference 166
FT /label= Gln, Arg
XX
XX WO9715667-A1.
XX
XX 01-MAY-1997.
XX
XX 25-OCT-1996; 96WO-US17201.
XX
XX 25-OCT-1995; 95US-0007015.
XX
XX (REG- ) REGENERON PHARM INC.
XX
XX Davis S, Gale NW, Yancopoulos GD;
XX
XX WPI: 1997-259021/23.
XX
XX N-PSDB; AAT69808.
XX
XX New nucleic acid encoding Efl-6 ligand protein - used for promoting
XX growth and proliferation of neuronal cells and in drug screening
XX
XX Claim 2; Fig 1; 36pp; English.
XX
XX A novel ligand (AAM17081), designated Efl-6 (or Eph transmembrane
XX tyrosine kinase family ligand 6), binds to the Elk, Nuk/Cek5,
XX HeK2/Seq4, Htk and Sek1 receptors on cells. Its amino acid
XX sequence was deduced from a human frontal cortex cDNA clone
XX (AAT69808). Recombinant Efl-6, truncated soluble polypeptides
XX comprising the extracellular domain of Efl-6, and Efl-6
XX ligand bodies comprising soluble Efl-6 and the FC portion of IgG can
XX be expressed in host cells. These can be used to support neuronal
XX and other Eph receptor-bearing cell populations for treatment of
XX neurological disorders, in drug screening and to raise diagnostic
XX antibodies.
XX
XX Sequence 340 AA;
SQ
Query Match 99.7%; Score 1844; DB 18; Length 340;
Best Local Similarity 99.7%; Pred. No. 9e-143;
Matches 339; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGPHSGPGGVRYGALLLGLVGLVSGLEPEYWNMSANKRFOAEGVLYPQIDRLDL 60
DB 1 MGPHSGPGGVRYGALLLGLVGLVSGLEPEYWNMSANKRFOAEGVLYPQIDRLDL 60
QY 61 LCPRAPPPGPHSSPNVEFYKLYVGAQGRCAAPPAPNLLTCRPPDLRLFTTKFOEY 120
DB 61 LCPRAPPPGPHSSPNVEFYKLYVGAQGRCAAPPAPNLLTCRPPDLRLFTTKFOEY 120
QY 121 SPFLMGHEPFSHDDYIITSDGTREGLSLOGVCLTREGMKVLLRVGSPRGCAVPRXP 180
DB 121 SPFLMGHEPFSHDDYIITSDGTREGLSLOGVCLTREGMKVLLRVGSPRGCAVPRXP 180
QY 121 SPFLMGHEPFSHDDYIITSDGTREGLSLOGVCLTREGMKVLLRVGSPRGCAVPRXP 180
DB 121 SPFLMGHEPFSHDDYIITSDGTREGLSLOGVCLTREGMKVLLRVGSPRGCAVPRXP 180
QY 181 VSMEMERDRGAASLEPEKENIPDPTSNATSRGAEGPLPPSPNPAVGAAGLALLL 240
DB 181 VSMEMERDRGAASLEPEKENIPDPTSNATSRGAEGPLPPSPNPAVGAAGLALLL 240
QY 241 GVAGAGAGAMCMRRRAKPSRSRHPGSGFRGSGSLGAGGAGPEAPRGELIALRGG 300
DB 241 GVAGAGAGAMCMRRRAKPSRSRHPGSGFRGSGSLGAGGAGPEAPRGELIALRGG 300
QY 301 GAADPPFCFHYEKVSGDYGHVYIVDDGPPQSPNNIYKYV 340
DB 301 GAADPPFCFHYEKVSGDYGHVYIVDDGPPQSPNNIYKYV 340

```

```

AC AAW33698;
XX 30-APR-1998 (first entry)
XX
XX AL-2-long (AL-21) protein.
XX
XX AL-2; AL-2; AL-2-long; human; treatment; neurological disorder; tumour;
XX rheumatoid arthritis; wound healing; paralysis; angiogenesis; leukaemia;
XX psoriasis; Alzheimer's disease; epilepsy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..26 "signal peptide"
XX Protein 27..455 "note= 'signal peptide'"
XX Domain 27..219 "note= 'mature protein'"
XX Domain 27..219 "note= 'mature protein'"
XX Domain 220..245 "note= 'extracellular domain'"
XX Domain 220..245 "note= 'extracellular domain'"
XX
XX WO9740153-A1.
XX
XX 30-OCT-1997.
XX
XX 17-APR-1997; 97WO-US06345.
XX
XX 19-APR-1996; 96US-0635130.
XX
XX (GETH ) GENENTECH INC.
XX
XX Caras IW;
XX
XX WPI: 1997-535837/49.
XX
XX N-PSDB; AAV06354.
XX
XX Human AL-2 neurotrophic factor and related DNA - used to develop
XX products for, e.g. treating neurologic disorder, angiogenesis
XX disorders, tumours or rheumatoid arthritis or for wound healing
XX
XX Claim 20; Fig 1A-C; 86pp; English.
XX
XX This is a AL-2-long (AL-21) protein. The AL-2 is a novel Eph-related
XX tyrosine kinase receptor ligand. AL-2 can be administered to patients in
XX whom the nervous system has been damaged by trauma, surgery, stroke,
XX ischaemia, infection, metabolic disease, nutritional deficiency,
XX malignancy, or toxic agents, to promote the survival or growth of
XX neurons. They can be used to treat motor neuron disorders such as
XX amyotrophic lateral sclerosis (Lou Gehrig's disease), Bell's palsy, and
XX various conditions involving spinal muscular atrophy, or paralysis. AL-2
XX can be used to treat human neurodegenerative disorders, such as
XX Alzheimer's disease, Parkinson's disease, epilepsy, demyelinating
XX diseases such as multiple sclerosis, Huntingtons chorea, Down's syndrome,
XX nerve deafness, Menier's disease, and other disorders of the cerebellum.
XX AL-2 can be used as cognitive enhancer, to enhance learning particularly
XX in dementia or trauma, since they can promote axonal outgrowth and
XX synaptic plasticity, particularly of hippocampal neurons that express
XX AL-2 binding Eph-family receptors and cortical neurons that express
XX AL-2. AL-2 can also be used for wound healing, i.e. accelerating
XX neovascularisation of, e.g. burns and ulcers. The encoding nucleic acids
XX are useful in preparing antibodies that specifically bind to the AL-2
XX protein. The antibodies and the AL-2 antagonists are useful in diagnosing
XX and treating various neuronal disorders. AL-2 antagonists can be used
XX for modulating angiogenesis. They can also be used for the treatment of
XX tumours, acute myeloid leukaemia (AML), chronic myeloid leukaemia (CML),
XX myelodysplastic syndrome (MDS), diabetic retinopathy, neovascular
XX glaucoma, psoriasis and rheumatoid arthritis.
XX
XX Sequence 455 AA;
SQ
Query Match 99.5%; Score 1841; DB 18; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.2e-142;

```

Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MGPHSGGGRVAGALLLGLVLSGLSLPEPTWNSANKRFOAEGGYLYPQIGRDL 60
DB 1 MGPHSGGGRVAGALLLGLVLSGLSLPEPTWNSANKRFOAEGGYLYPQIGRDL 60
QY 61 LCPRARPPGPSSPNVEFYKLYLVGAQGRCEAPAPNLLTCDRPDLDFRTIKFOEY 120
DB 61 LCPRARPPGPSSPNVEFYKLYLVGAQGRCEAPAPNLLTCDRPDLDFRTIKFOEY 120
QY 121 SPNMGHEFRSHHDYIYIATSDGTREGLESIOGGVCLTRGMKVLIRVQSPRGAVPRKP 180
DB 121 SPNMGHEFRSHHDYIYIATSDGTREGLESIOGGVCLTRGMKVLIRVQSPRGAVPRKP 180
QY 181 VSEMPMERDRGAASLEPGKENTLPGDPTSNATSGAAGPLPPSPMPAVAGAAGLALLL 240
DB 181 VSEMPMERDRGAASLEPGKENTLPGDPTSNATSGAAGPLPPSPMPAVAGAAGLALLL 240
QY 241 GVAAGAGAMCWRRRRAKPSRRHPGSGFGRGSLGIGGGGGMGPRAEPGELGIALRG 300
DB 241 GVAAGAGAMCWRRRRAKPSRRHPGSGFGRGSLGIGGGGGMGPRAEPGELGIALRG 300
QY 301 GAADPPFCPHYEKVSGDYGHPIYIVDGPQSPNITY 338
DB 301 GAADPPFCPHYEKVSGDYGHPIYIVDGPQSPNITY 338

```

RESULT 6

AAW4615
ID AAW4615 standard; Protein; 340 AA.

AC AAW4615;
XX
DT 06-JUL-1998 (first entry)
XX

DE Human transmembrane ligand Elk-L3.

XX Elk-L3; Elk-related receptor tyrosine kinase; transmembrane ligand;
KW human; signal transduction; axonogenesis; nerve cell; neurone;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW demyelination; multiple sclerosis; amyotrophic lateral sclerosis;
KW nervous system infection; Wernicke's disease; trauma; ischaemia;
KW stroke; nutritional polyneuropathy; progressive supranuclear palsy;
KW Shy Drager's syndrome; multistep degeneration;
KW olivoponto cerebellar atrophy; peripheral nerve damage.

OS Homo sapiens.

XX
FH Key Location/Qualifiers

FT Misc-difference 166
FT /label= Gln, Arg
FT 225..249

FT Domain /note="transmembrane domain"

PN MO9801548-A1.

XX 15-JAN-1998.

XX 04-JUL-1997; 97MO-CA00473.

XX 05-JUL-1996; 96US-0021272.

XX (MOUN) MOUNT SINAI HOSPITAL CORP.

XX PA Holland S, Mdamaju G, Pawson T;

XX WPI; 1998-101047/09.

XX DR N-PSDB; AAV16097.

XX
PT Modulating transmembrane ligand for an Elk-related receptor tyrosine
kinase - by formation of a complex between an oligomerised
PT Elk-related receptor tyrosine kinase and a transmembrane ligand
XX

PS Disclosure; Fig 5A; 40pp; English.

XX This polypeptide comprises human Elk-L3, a transmembrane ligand of
CC Elk-related receptor tyrosine kinase (ERK). A novel method of
CC modulating the biological activity of, or for affecting a pathway
CC regulated by, a transmembrane ligand for an ERK in a cell
CC expressing the transmembrane ligand comprises forming a complex
CC between a purified and isolated oligomerised ERK, or an isoform
CC or an extracellular domain of the ERK, and the transmembrane
CC ligand expressed on the cell. The complex can also be used for
CC evaluating a substance for its ability to modulate the biological
CC activity of a transmembrane ligand for an ERK, and to identify
CC substances that affect or modulate a pathway regulated by a ERK.
CC A purified and isolated oligomerised ERK can be used in the
CC preparation of a medicament for modulating neuronal development or
CC regeneration in a subject, or in a medicament for modulating
CC axonogenesis in a subject (all claimed). The substances identified
CC by the methods can be used to modulate axonogenesis, nerve cell
CC interactions and regeneration, to treat diseases and conditions
CC involving trauma and injury to the nervous system, such as
CC Alzheimer's disease, Parkinson's disease, Huntington's disease,
CC demyelinating diseases, such as multiple sclerosis, amyotrophic
CC lateral sclerosis, bacterial and viral infections of the nervous
CC system, deficiency diseases, such as Wernicke's disease and
CC nutritional polyneuropathy, progressive supranuclear palsy,
CC Shy Drager's syndrome, multistep degeneration and olivoponto
CC cerebellar atrophy, peripheral nerve damage, trauma, and ischaemia
CC resulting from stroke.

XX Sequence 340 AA;

Query Match 99.2%; Score 1836; DB 19; Length 340;
Best Local Similarity 99.4%; Pred. No. 4e-142;
Matches 338; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MGPHSGGGRVAGALLLGLVLSGLSLPEPTWNSANKRFOAEGGYLYPQIGRDL 60
DB 1 MGPHSGGGRVAGALLLGLVLSGLSLPEPTWNSANKRFOAEGGYLYPQIGRDL 60
QY 61 LCPRARPPGPSSPNVEFYKLYLVGAQGRCEAPAPNLLTCDRPDLDFRTIKFOEY 120
DB 61 LCPRARPPGPSSPNVEFYKLYLVGAQGRCEAPAPNLLTCDRPDLDFRTIKFOEY 120
QY 121 SPNMGHEFRSHHDYIYIATSDGTREGLESIOGGVCLTRGMKVLIRVQSPRGAVPRKP 180
DB 121 SPNMGHEFRSHHDYIYIATSDGTREGLESIOGGVCLTRGMKVLIRVQSPRGAVPRKP 180
QY 181 VSEMPMERDRGAASLEPGKENTLPGDPTSNATSGAAGPLPPSPMPAVAGAAGLALLL 240
DB 181 VSEMPMERDRGAASLEPGKENTLPGDPTSNATSGAAGPLPPSPMPAVAGAAGLALLL 240
QY 241 GVAAGAGAMCWRRRRAKPSRRHPGSGFGRGSLGIGGGGGMGPRAEPGELGIALRG 300
DB 241 GVAAGAGAMCWRRRRAKPSRRHPGSGFGRGSLGIGGGGGMGPRAEPGELGIALRG 300
QY 301 GAADPPFCPHYEKVSGDYGHPIYIVDGPQSPNITY 340
DB 301 GAADPPFCPHYEKVSGDYGHPIYIVDGPQSPNITY 340

```

RESULT 7

AAW00287
ID AAW00287 standard; Protein; 334 AA.

XX AAW00287;

XX 19-JAN-1997 (first entry)

XX Mouse Eph receptor ligand ELF-2.

XX Eph receptor; ligand; ELF-2; tyrosine kinase; signal transduction;
KW axonogenesis; oncogenesis; tumour; neurological disorder;
KW diagnosis; gene therapy.

```

XX OS Mus sp.
XX FH Key
XX FT Domain
XX FT Binding-site
XX FT Domain
XX FT WO9626958-A2.
XX PN 06-SEP-1996.
XX PD 23-FEB-1996, 96WO-US02673.
XX PR 27-FEB-1995, 95US-0395415.
XX XX (HARD ) HARVARD COLLEGE.
XX PA Bergemann AD, Flanagan JG;
XX PI WPI; 1996-433391/43.
XX DR N-PSDB; AAT40230.
XX PS Claim 6; Fig 1A-B; 50pp; English.
XX CC Mouse Eph receptor ligand ERF-2 (AA00287) is strongly expressed in
XX CC the anterior hindbrain and newly-forming somites of embryos at the
XX CC early organogenesis stage of development. It is important in
XX CC cellular communication during pattern formation. Its amino acid
XX CC sequence was deduced from a cDNA clone (AAT40230) isolated from a
XX CC newborn mouse brain cDNA library. The ERF-2 ligand can be used
XX CC to alter neurological development, oncogenesis and growth
XX CC regulation, to modulate binding of ERF-2 to the Eph receptor, and
XX CC in diagnostic assays.
XX SQ Sequence 334 AA;
XX
XX Query Match 34.5%; Score 637.5; DB 17; Length 334;
XX Best Local Similarity 42.1%; Pred. No. 4,6e-44;
XX Matches 139; Conservative 49; Mismatches 129; Indels 13; Gaps 5;
XX
XX QY 14 GALLILGVLGVLSLEFVYNNNSANKRQAEQGYLYPIQIDRLDLCPRARPPGPHSS 73
XX DB 15 GLMLVLCRTAISRIVLEPIYNNSSNSKFLPGQGLVLPQIDKDLICPKV--DSKTV 71
XX QY 74 PNYEFKLYLVGAQGRCEAPPAFNLITCDRDLRFTIKFOEYSPNLTMGHEFRSHH 133
XX DB 72 GQYEVYKVMVDQADRCTIKKENTPLNLCARPDDQVFTIKFOEFSNLMGLEFOQNK 131
XX
XX QY 134 DYTIIATSDGTREGLSLOGVCLTGKMYLLRVGQ--SPRGAVPRKPVSEMPMR-DR 190
XX DB 132 DYTIIATSDGTREGLSLOGVCLTGKMYLLRVGQ--SPRGAVPRKPVSEMPMR-DR 191
XX QY 191 GAHSHLEPGKENVLPDPTNATSRGAEGLPPSPMAVAGAGLALLLGVAGAGANVC 250
XX DB 192 GRSTSTSPFKVPDSSGTDGNSAGHSGNLLGSEVALFPGISGCIIFVITLVLVL 251
XX QY 251 WRRRRARFSESRHPPGSGFGRGSLGCGGGKGPREAEPGELGIALRGGAADPPFCPH 310
XX DB 252 KYRRRRARKSPQHTTTLSTLATPRKGAN-----NGSEPSVITPLR---TADSVFCBH 304
XX QY 311 YEKVSGDYGHPIYIVODGPOSPNLYYKV 340
XX DB 305 YEKVSGDYGHPIYIVODGPOSPNLYYKV 334

```

```

XX RESULT 8
XX AAR92742
XX ID AAR92742 standard; Protein; 336 AA.
XX XX
XX AC AAR92742;
XX XX
XX DT 21-MAY-1996 (first entry)
XX XX
XX DE Murine hepatoma transmembrane kinase receptor ligand.
XX XX
XX KM Hepatoma transmembrane kinase; Htk; receptor; ligand;
XX KM tyrosine kinase; neurodegenerative disease.
XX XX
XX OS Mus musculus.
XX XX
XX PN WO9602645-A2.
XX PD 01-FEB-1996.
XX XX
XX PF 14-JUL-1995, 95WO-US08812.
XX XX
XX PR 20-JUL-1994, 94US-0277722.
XX XX
XX PA (GETH ) GENENTECH INC.
XX XX
XX PI Bennett BD, Matthews W;
XX XX
XX DR WPI; 1996-105907/11.
XX DR N-PSDB; AAT16470.
XX XX
XX PT Ligand for the hepatoma trans-membrane kinase receptor - useful for
XX PT stimulating and inhibiting cells carrying the receptor, e.g. for
XX PT treating neuro-degenerative disease
XX XX
XX PS Claim 5; Fig 1(A-D); 86pp; English.
XX XX
XX CC Mouse (AAT16470) and human (AAT16471) Htk ligand which bind to, and
XX CC activate, the Htk receptor, have been identified in a variety of
XX CC tissues using a soluble Htk-Fc fusion protein.
XX CC The predicted mol.wt. of the murine Htk ligand protein following a
XX CC signal peptide cleavage is 34 kD with an estimated pI of 8.9.
XX CC The murine and human ligands show 96% homology at the amino acid
XX CC level.
XX CC The DNA is used to produce recombinant ligands; for tissue-
XX CC specific typing (partic. as a marker for breast cancer) and as a
XX CC marker for human chromosome 13. The ligands (partic. in soluble
XX CC form) are used to activate the tyrosine kinase domain of the
XX CC Htk receptor, i.e. to stimulate or inhibit growth, differentiation,
XX CC and/or activation of cells contg. the receptor, e.g. treatment
XX CC of neurodegenerative diseases, since they are strongly expressed
XX CC in the cerebral cortex, hippocampus, striatum and cerebellum.
XX CC The ligands are also useful as a control or standard in assays,
XX CC for generation of antibodies, as a mol. wt. marker, for growth
XX CC in vitro of Htk-receptor positive cells, as research agent,
XX CC in screening, etc.
XX XX
XX SQ Sequence 336 AA;
XX
XX Query Match 34.5%; Score 637.5; DB 17; Length 336;
XX Best Local Similarity 42.1%; Pred. No. 4,6e-44;
XX Matches 139; Conservative 49; Mismatches 129; Indels 13; Gaps 5;
XX
XX QY 14 GALLILGVLGVLSLEFVYNNNSANKRQAEQGYLYPIQIDRLDLCPRARPPGPHSS 73
XX DB 17 GLMLVLCRTAISRIVLEPIYNNSSNSKFLPGQGLVLPQIDKDLICPKV--DSKTV 73
XX QY 74 PNYEFKLYLVGAQGRCEAPPAFNLITCDRDLRFTIKFOEYSPNLTMGHEFRSHH 133
XX DB 74 GQYEVYKVMVDQADRCTIKKENTPLNLCARPDDQVFTIKFOEFSNLMGLEFOQNK 133
XX QY 134 DYTIIATSDGTREGLSLOGVCLTGKMYLLRVGQ--SPRGAVPRKPVSEMPMR-DR 190
XX DB 134 DYTIIATSDGTREGLSLOGVCLTGKMYLLRVGQ--SPRGAVPRKPVSEMPMR-DR 193

```

```

QY 191 GAHSLERKENTLPGDPTSNATSRGAEGLPPSPMPAVAGAAGLALLLGVAGAGAMC 250
DB 194 GRSTTPEFVKNPFGSSITDGNASHGNNLLGSEVALPAIAGGCIIFVITITLVLL 253
QY 251 WRARRAPSESRHHPGSGFGRGSLGCGGCGMPREAREBELGIALRGGAADPPFCPH 310
DB 254 KYRRRHRKHSPQHTTTLSTLSTLATPKRGNN---NGSEPSDVIPLR---TDSVFCPH 306
QY 311 YEKSGDYGHPVYIVQGPQSPNNITYKV 340
DB 307 YEKSGDYGHPVYIVQEMPQSPANNITYKV 336

RESULT 9
AAR55059 standard; Protein; 346 AA.
AC AAR55059;
XX 28-JAN-1995 (first entry)
XX
DE Elk tyrosine kinase receptor ligand.
XX Vectors; elk-L protein; elk; ligands; cell growth; differentiation.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Peptide 1..24
FT Protein /note="signal peptide"
FT Protein /note="mature elk-L protein"
XX
XX NO9411384-A.
XX
XX 26-MAY-1994.
XX
XX 15-NOV-1993; 93MO-US10955.
XX
XX 13-NOV-1992; 92US-0977693.
XX
XX (IMMV ) IMMUNEX CORP.
XX
XX Baum PR, Beckmann MP, Lyman S;
XX
XX WPI; 1994-183415/22.
XX
XX N-PSDB; AAG65486.
XX
XX New DNA encoding ligand for elk tyrosine kinase receptor - also
XX related polypeptides, vectors, antibodies and probes, useful e.g.
XX in studying cell differentiation or growth
XX
XX Claim 7; Page 30; 35pp; English.
XX
XX The sequence is that of the elk-L protein able to bind elk, a
XX tyrosine kinase receptor. The DNA may be incorporated into vectors
XX which can be used to study the role of elk and its ligands in cell
XX growth and differentiation.
XX
XX Sequence 346 AA;

Query Match 34.2%; Score 632; DB 15; Length 346;
Best Local Similarity 39.5%; Pred. No. 1.3e-43;
Matches 145; Conservative 48; Mismatches 116; Indels 58; Gaps 9;
QY 8 PGGRVGAALLLGLVLSGL-----SLRPVWNSANKRFOAGGYVYPOIGDRIDL 61
DB 4 PGGRWLKXMLVAVVWMLCRLATPLAKNLEFVWSSINRKLVSGLVYIPKIGKDLII 63
QY 62 CPDAPPPGPHSSPNVEFYKYLIVGAGGRRCEAPPAVNLITCDRPLDLRFTIKQEYS 121
DB 64 CPDAPPPGPHSSPNVEFYKYLIVGAGGRRCEAPPAVNLITCDRPLDLRFTIKQEYS 121
DB 64 CPDAPPPGPHSSPNVEFYKYLIVGAGGRRCEAPPAVNLITCDRPLDLRFTIKQEYS 121

```

```

QY 122 PNLNGHPRSHNDYITITSDGTREGLSLGGVCLITGKAVLLRVGSGPRGAVPRKV 181
DB 119 PNYMLEFERKHHDDYITITSDGSLGLENRREGVCRITNMKIINKVQDDPAVAVPEOLTT 178
QY 182 SEMPERRGAHSLR-GKENLPGDPTSNATSRGAEGLPPSPMPAVAGAAGLA---- 236
DB 179 SPSSEADANTYMATQAGSGRSLGDSGCKHETYNQEKSP-----GASGSSGDPD 231
QY 237 -----LILLGVAGAGA-----MCWRRRAPSESRHHPGSGFGRGSLGL 277
DB 232 GFENSKVALFAVAGGCVIFLLIIFLTVLLKLRKRHKHTQO-----RAAALSL 282
QY 278 ---GGGCGMPREAREBELGIALRGGAADPPCPHYEKVSGYGHVYIVQGPQSP 333
DB 283 STLAPKGGSGTAGTEPBDITILPLR---TTENNYCPHYEKVSGYGHVYIVQEMPQSP 339
QY 334 PNITYKV 340
DB 340 ANITYKV 346

RESULT 10
AAR91930 standard; Protein; 346 AA.
AC AAR91930;
XX 11-DEC-1996 (first entry)
XX
XX AAR91930;
XX
XX Key Location/Qualifiers
FT Peptide 1..24
FT Peptide /label= sig_peptide
FT Peptide 25..346
FT Peptide /label= mat_peptide
XX
XX US5512457-A.
XX
XX 30-APR-1996.
XX
XX 13-NOV-1992; 92US-0977693.
XX
XX 15-MAR-1994; 94US-0213403.
XX
XX 13-NOV-1992; 92US-0977693.
XX
XX (IMMV ) IMMUNEX CORP.
XX
XX Baum PR, Beckmann MP, Carpenter MK, Lyman S;
XX
XX WPI; 1996-229866/23.
XX
XX N-PSDB; AAT28770.
XX
XX DNA coding for neurotrophic human elk ligand cytokine - useful as
XX probe to isolate other elk ligand sequences
XX
XX Claim 1; Columns 27-30; 18pp; English.
XX
XX The present sequence is the human cytokine elk-ligand (elk-L),
XX which binds a member of the tyrosine kinase receptor family. Elk-L
XX exhibits neurotrophic and neuroprotective properties, and has a
XX calculated mol. wt. 35180 daltons and a pI of 9.006. The elk-L
XX cDNA, isolated from a human placental cDNA library, can be
XX radiolabelled and used as a probe for isolating other mammalian
XX elk-L cDNA. Elk-L can be used to treat neural tissue disorders,

```

CC partic. excito-toxicity associated injuries or disorders, and as a
 CC neural culture reagent, while immunogenic fragments of elk-L can be
 CC used to generate specific anti-elk-L antibodies.

CC Sequence 346 AA;

Query Match 34.2%; Score 632; DB 17; Length 346;
 Best Local Similarity 39.5%; Pred. No. 1.3e-43;
 Matches 145; Conservative 48; Mismatches 116; Indels 58; Gaps 9;

QY 8 PGGVVGAALLLVGLVLSGL-----SLSPYVNSANKRFOAGGYLYLPQIGRLDL 61
 DB 4 PGGRWLGKMLVAVVWVALCRLATPLAKNLEPVNSGSLNPKFSGKGLVYIPKIGDLII 63
 QY 62 CPBARPPGHSBNPEFYKLYLVGAQGRCEAPAPNLLTCDPDLRLFTIKFOEYS 121
 DB 64 CPBARAGRP-----YEVYKLYVREPOAAACSTVLDPNVLTVCNREDEIRFTIKFOEFS 118
 QY 122 PNLMGHEFRSHHDYIATSDGTREGLESLOGVCLTRGMKYLRLVQSGPRGGAVERKPV 181
 DB 119 PNIMGLEFKKHHDYITTSNSGLEENREGVCRTKIMKVGODPNAVTPQULTT 178
 QY 182 SEMPMERDRGAASLE-PGKENVLPDPTSNATSRGAEGLPPSPMAVAGAAGLA---- 236
 DB 179 SRPSKADNTVMKATQAPSGRSLGSDGKHETVNOEKSQP-----GASGSSGDPD 231
 QY 237 -----LILLGVAGAGA-----MCMRRRAKPSBSRHPGSGRGSLGL 277
 DB 232 GFNSKVALFAVAGACVIFLLIIFLVLLKLRKHKHQ-----RAAALSL 282
 QY 278 -----GGGGGMPREAPRGELGIALRGGAADPPFCPHYEKVSGDYHPVYIVODGPQSP 333
 DB 283 STLASPRGSGTAGTEPSDIIIPLR---TTENNYCPHYEKVSGDYGHVYIVQENMPQSP 339
 QY 334 PNITYYKV 340
 DB 340 ANITYYKV 346

RESULT 11

AAW36055 standard; Protein; 346 AA.

AC AAW36055;
 XX 06-MAR-1998 (first entry)
 DT Human elk-L protein.
 XX Human elk-L protein.
 KW Human, elk-L, cytokine, ligand, tyrosine kinase receptor, fusion protein;
 KW extracellular domain, immunoglobulin, neurological disease.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Peptide 1..24
 FT /note= "signal peptide"
 FT Protein 25..346
 FT /note= "mature protein"
 FT Domain 25..237
 FT /note= "extracellular domain; this region is used to
 generate a fusion protein with the Fc portion
 of the human immunoglobulin G1"
 FT Domain 238..265
 FT /note= "transmembrane domain"
 FT Domain 266..346
 FT /note= "intracellular domain"
 FT Modified-site 139..141
 FT /note= "Asn is N-glycosylated"
 FT Cleavage-site 266..267
 FT /note= "KEX2 protease cleavage site"
 FT Cleavage-site 267..268
 FT /note= "KEX2 protease cleavage site"

FT Cleavage-site 270..271
 FT /note= "KEX2 protease cleavage site"
 XX US5670625-A.
 XX PD 23-SEP-1997.
 XX 02-JUN-1995; 95US-0460741.
 XX 15-MAR-1994; 94US-0213403.
 XX 13-NOV-1992; 92US-0977693.
 XX 02-JUN-1995; 95US-0460741.
 XX (IMMEX CORP.
 XX Baum PR, Beckmann MP, Lyman S;
 XX WPI; 1997-479524/44.
 DR N-PADB; AAT97976.
 PT Soluble fusion proteins of human elk-L ligand and Fc immunoglobulin
 PT fragment - and their dimers and oligomers, useful as
 PT neuro-protectants and neurotrophic agents
 XX Claim 1; Columns 27-30; 18pp; English.

CC This is the amino acid sequence of the human elk-L protein, a new
 CC cytokine that is the ligand for the elk tyrosine kinase receptor. The
 CC extracellular domain of the protein (amino acids 1-213) is used to
 CC generate a fusion protein comprising the Fc polypeptide of the human
 CC immunoglobulin G1 (extending from the hinge region to the C-terminus).
 CC The fusion protein (which has the same activities as the natural elk-L
 CC protein) has neuroprotective and neurotrophic activity so is potentially
 CC useful for treating a wide range of neurological diseases.

XX Sequence 346 AA;

Query Match 34.2%; Score 632; DB 18; Length 346;
 Best Local Similarity 39.5%; Pred. No. 1.3e-43;
 Matches 145; Conservative 48; Mismatches 116; Indels 58; Gaps 9;

QY 8 PGGVVGAALLLVGLVLSGL-----SLSPYVNSANKRFOAGGYLYLPQIGRLDL 61
 DB 4 PGGRWLGKMLVAVVWVALCRLATPLAKNLEPVNSGSLNPKFSGKGLVYIPKIGDLII 63
 QY 62 CPBARPPGHSBNPEFYKLYLVGAQGRCEAPAPNLLTCDPDLRLFTIKFOEYS 121
 DB 64 CPBARAGRP-----YEVYKLYVREPOAAACSTVLDPNVLTVCNREDEIRFTIKFOEFS 118
 QY 122 PNLMGHEFRSHHDYIATSDGTREGLESLOGVCLTRGMKYLRLVQSGPRGGAVERKPV 181
 DB 119 PNIMGLEFKKHHDYITTSNSGLEENREGVCRTKIMKVGODPNAVTPQULTT 178
 QY 182 SEMPMERDRGAASLE-PGKENVLPDPTSNATSRGAEGLPPSPMAVAGAAGLA---- 236
 DB 179 SRPSKADNTVMKATQAPSGRSLGSDGKHETVNOEKSQP-----GASGSSGDPD 231
 QY 237 -----LILLGVAGAGA-----MCMRRRAKPSBSRHPGSGRGSLGL 277
 DB 232 GFNSKVALFAVAGACVIFLLIIFLVLLKLRKHKHQ-----RAAALSL 282
 QY 278 -----GGGGGMPREAPRGELGIALRGGAADPPFCPHYEKVSGDYHPVYIVODGPQSP 333
 DB 283 STLASPRGSGTAGTEPSDIIIPLR---TTENNYCPHYEKVSGDYGHVYIVQENMPQSP 339
 QY 334 PNITYYKV 340
 DB 340 ANITYYKV 346

RESULT 12

AAW19249 standard; Protein; 346 AA.

```

XX AC AAW19249;
XX 18-AUG-1997 (first entry)
XX DT
XX XX Human elk ligand protein.
XX XX
XX XX Human; elk; ligand; elk-L; cytokine; testing; measuring;
XX KM purification; neuroprotection; treatment; diabetic; hereditary;
XX KM nutritional; neuropathy; neurodegenerative disease;
XX XX tissue culture.
XX OS Homo sapiens.
XX XX
XX XX Key Location/Qualifiers
XX FT Peptide 1..24
XX FT /label= sig_peptide
XX FT Peptide 25..346
XX FT /label= mat_peptide
XX FT
XX PN US5627267-A.
XX XX
XX PD 06-MAY-1997.
XX PF 13-NOV-1992; 92US-0977693.
XX XX
XX PR 15-MAR-1994; 94US-0213403.
XX PR 13-NOV-1992; 92US-0977693.
XX PR 01-JUN-1995; 95US-0458077.
XX XX
XX PA (IMMUNEX CORP.
XX XX
XX PI Baum PR, Beckmann MP, Lyman S;
XX XX
XX DR WPI; 1997-271366/24.
XX DR N-PSDB; AAT69766.
XX XX
XX PT Human elk ligand protein - for diagnostic or therapeutic use, e.g.
XX FT as neuro-protective agent
XX XX
XX PS Claim 1; Columns 29-32; 18pp; English.
XX XX
XX CC The present sequence is a human elk ligand (elk-L) protein,
XX CC which binds elk, has a calculated molecular weight of 35180 and an
XX CC isoelectric point of 9.006. Elk-L is a cytokine that can be used to
XX CC test cells for elk expression, measure the biological activity of
XX CC elk, purify elk by affinity chromatography and as a neuroprotective
XX CC agent to treat diabetic, hereditary and nutritional neuropathies
XX CC and neurodegenerative diseases. It may also be added to tissue
XX CC cultures to prolong the life of neurons. The elk-L cDNA was
XX CC isolated from a human placental cDNA library, and is present as a
XX CC cDNA insert in the recombinant vector deposited in strain
XX CC ATCC 69085.
XX XX
XX SQ Sequence 346 AA;

```

Query Match 34.2%; Score 632; DB 18; Length 346;
 Best Local Similarity 39.5%; Pred. No. 1.3e-43;
 Matches 145; Conservative 48; Mismatches 116; Indels 58; Gaps 9;

```

DB 179 SRPSKEADNTVMATQARSGSLGSDGKHETVVOEKEGP-----GASGSSGSPD 231
QY 237 -----LILIGVAGAGA-----MCMRRRAKPSERHPGSGFRGSGISGL 277
DB 232 GFENSKVALFAVAGACVIFLLIIFLTVLLIKRKRRKHTQQ-----RAAALSL 282
QY 278 ----GGGGMGPREAPPELIGIALRGGAADPPCPHYEKVSGYGHVYIVODGPPSP 333
DB 283 SLLAPKSGSGTACGPEPDITIPLR--TENNVCPIREKVSGLGHFVYIVQENPPSP 339
QY 334 PNIIYKV 340
DB 340 ANIYKV 346

```

RESULT 13
 AAW44323
 ID AAW44323 standard; Protein; 346 AA.
 XX AC AAW44323;
 XX XX
 XX DT 27-MAY-1998 (first entry)
 XX XX
 XX DE Human elk-L.
 XX XX
 XX KM Human; elk-L; elk ligand; cytokine; antibody; cell surface receptor;
 XX KM tyrosine kinase receptor; neural disease; trophic mechanism.
 XX OS Homo sapiens.
 XX XX
 XX XX Key Location/Qualifiers
 XX FT Peptide 1..24
 XX FT /label= signal
 XX FT Protein 25..346
 XX FT /label= elk-L
 XX FT
 XX PN US5728813-A.
 XX XX
 XX PD 17-MAR-1998.
 XX PF 12-NOV-1996; 96US-0747240.
 XX XX
 XX PR 15-MAR-1994; 94US-0213403.
 XX PR 13-NOV-1992; 92US-0977693.
 XX PR 02-JUN-1995; 95US-0460741.
 XX PR 12-NOV-1996; 96US-0747240.
 XX XX
 XX PA (IMMUNEX CORP.
 XX XX
 XX PI Baum PR, Beckmann MP, Lyman S;
 XX XX
 XX DR WPI; 1998-206621/18.
 XX DR N-PSDB; AAV15226.
 XX XX
 XX PT Antibodies to elk ligand polypeptides - bind to tyrosine kinase
 XX FT receptors; useful for treating neural disease
 XX XX
 XX PS Claim 1; Column 31-32; 17pp; English.
 XX XX
 XX CC The present sequence represents human elk-L (elk ligand). The present
 XX CC invention describes antibodies to elk-L polypeptides. The elk-L protein
 XX CC is known to be neuroprotective exhibiting its effect through a trophic
 XX CC mechanism. Examples of diseases that may be treated with elk-L include
 XX CC neuropathies e.g. diabetic, hereditary and nutritional neuropathies,
 XX CC neurodegenerative diseases and other diseases characterised by loss of
 XX CC function or degeneration of neurons.
 XX XX
 XX SQ Sequence 346 AA;

Query Match 34.2%; Score 632; DB 19; Length 346;
 Best Local Similarity 39.5%; Pred. No. 1.3e-43;
 Matches 145; Conservative 48; Mismatches 116; Indels 58; Gaps 9;

Wed Feb 12 09:30:13 2003

Db 304 YEKVSGDYGHFVYIVQEMPPQSPANIYYKV 333

RESULT 19

AA89287

AA89287 standard; Protein; 333 AA.

AC AAR89287;

XX 13-APR-1996 (first entry)

DT Human LERK-5 cytokine.

XX Human LERK-5 cytokine; elk; hek; protein-tyrosine-kinase;

XX Human; LERK-5; cytokine; elk; hek; protein-tyrosine-kinase;

XX receptor; chromosome-13; expressed sequence tag; T-lymphocyte;

XX fetal brain; cloning; soluble LERK-5; antibody; Fc; fusion protein;

XX dimer; FLAG; affinity tail; drug delivery; cell targeting;

XX cytostatic; leukaemia; therapy.

XX Homo sapiens.

OS

XX

XX Key

XX Peptide

FT 1..25 Location/Qualifiers

FT 1..224 "Signal peptide"

FT 1..224 "Soluble LERK-5 (claim 3)"

FT 26..37 "N-terminal peptide (claim 15)"

FT 26..224 "Extracellular domain (mature soluble LERK-5, claim 3)"

FT Modified-site

FT 36..38 /note= "N-glycosylation site"

FT 62 /note= "Conserved cysteine residue"

FT 89 /note= "Conserved cysteine residue"

FT 101 /note= "Conserved cysteine residue"

FT 104..215 /note= "Chromosome-13 expressed sequence tag Genbank L13189 translation"

FT Modified-site

FT 139..141 /note= "N-glycosylation site"

FT 153 /note= "Conserved cysteine residue"

FT 225..250 /note= "Transmembrane region"

FT 251..333 /note= "Cytoplasmic domain"

FT 253..254 /note= "KEX2 protease processing site"

FT 254..255 /note= "KEX2 protease processing site"

FT 257..258 /note= "KEX2 protease processing site"

FT 276..277 /note= "KEX2 protease processing site"

FT /note= "KEX2 protease processing site"

XX WO9601839-A1.

XX 25-JAN-1996.

XX

XX 06-JUL-1995; 95MO-US08534.

XX 08-JUL-1994; 94US-0271948.

XX (IMMUNEX CORP.

XX Cerretti DP, Reddy P;

XX WPI; 1996-087585/10.

XX N-PSDB; AA099288.

XX DNA encoding LERK-5 cytokine capable of binding elk and hek - used

PT as carrier to deliver attached (anti-leukaemic) agents to cells

PT bearing elk or hek, i.e. human leukaemia cell lines JM and LK63

XX

PS Claim 1; Page 31-32; 40pp; English.

XX

CC The sequence represents a full-length LERK-5 cytokine, which binds

CC elk and hek receptor protein-tyrosine-kinases (members of the

CC eph/elk family), and expressed from a gene isolated by amplification

CC with primers specific for a chromosome-13 expressed sequence tag,

CC Genbank L13189, from a human T-lymphocyte leukaemia CCRF-HSB-2 (ATCC

CC CCL 120.1) cell line, followed by cDNA probe screening of a human

CC fetal brain cDNA library in phage lambda-gt10. A truncated soluble

CC form may also be produced. The protein may be produced as a

CC recombinant fusion protein or dimer with an antibody Fc fragment or

CC e.g. a FLAG affinity tail (R89285). LERK-5 may be used in binding

CC assays to detect elk- or hek-expressing cells, to measure elk or hek

CC activity, as a drug delivery for elk or hek receptor-specific cell

CC targeting, e.g. for delivery of cytostatic agents in leukaemia

CC therapy.

CC

XX

XX Sequence 333 AA;

XX

XX Query Match 34.0%; Score 629.5; DB 17; Length 333;

XX Best Local Similarity 40.9%; Pred. No. 2.1e-43;

XX Matches 135; Conservative 52; Mismatches 130; Indels 13; Gaps 5;

QY 14 GALLLLGVLGLVSGLSLEPVYVNSANKRFQAGGVLYTPQIGDLDLDCPPARPPGHSS 73

Db 14 GVLWVLCRTAISKSVILEPIYVNSNSKFLPGGLVLYPQIGDLDLDCPKV--DSKTV 70

QY 74 PNYEYKLYLVGAQGRCEAPAPNILLTCDPDLDFRTKFEYSFNLGHEFRSH 133

Db 71 GQYEVYKVMYDQKQARCTTKENTPLNCARQDQIKFTIKQESFNLGHEFRSH 130

QY 134 DYIYIATSDGTRGIESAQGVCLTRGKVLIRVGO--SPRGAVPRKPSSEMER-DR 190

Db 131 DYIYIISTNGSLLEGIDNDEGVCCYTRMKILMKVGQDASSAGSTRNDPTRPPELEGTN 190

QY 191 GAHSLBEGKKNLPQDDPFSNATSRGAEPLPPMPVAVAGAGLALLLVAGAGAMC 250

Db 191 GRSTTSPFVKNPNOSSSTDGNSAGHSNNILGSEVALPAGIASCCIFIVIIILVLL 250

QY 251 WRRRAKPSERHPQPSFGRGSLGLGGGGMGPRELPEGLAIALRGGAADPPCPH 310

Db 251 KYRRHRKHSPOHTTTLSLTATPKRSNN---NGSEPDIIIPUR---TADSVCPH 303

QY 311 YEKVSGDYGHFVYIVQDGPQSPANIYYKV 340

Db 304 YEKVSGDYGHFVYIVQEMPPQSPANIYYKV 333

XX

XX RESULT 20

XX AA11308

XX ID AA11308 standard; protein; 333 AA.

XX AC AA11308;

XX

XX 21-MAR-1997 (first entry)

XX

XX Receptor-type tyrosine kinase #4.

XX

XX Receptor-type tyrosine kinase; LPM medium; ligand; autophosphorylation;

XX insect cell; animal cell; growth promoter; undifferentiated cell;

XX insulin; interleukin; fibroblast growth factor; hepatocyte growth factor;

XX nerve growth factor; interferon-gamma; tumour necrosis factor; inducer.

XX Homo sapiens.

XX

XX Key

XX Peptide

FT 1..25 Location/Qualifiers

FT /label= signal peptide

```

FT Protein 26..333
FT /label= receptor-type tyrosine kinase
XX
XX JP08280385-A.
XX
XX 29-OCT-1996.
XX
XX 13-APR-1995; 95JP-0087878.
XX
XX 13-APR-1995; 95JP-0087878.
XX
XX (ASAH ) ASAH KASEI KOGYO KK.
XX
XX WPI; 1997-014848/02.
XX
XX N-PSDB; AAT51236.
XX
XX LPM cell culture medium - for culturing cells which produce
PT receptor-type tyrosine kinase ligand
XX
XX Disclousre; Page 40-41; 45pp; Japanese.
XX
XX AA011307 and AA011308 represent receptor-type tyrosine kinases. These
CC sequences are included in the LPM medium of the invention. The medium
CC is used to culture cells capable of producing a ligand for receptor-type
CC tyrosine kinase. The ligand serves as a physiologically active
CC substance to induce enzyme activity and autophosphorylation of tyrosine
CC kinase, which is involved in the development and differentiation of
CC animal and insect cells. The ligand can also be used as in a
CC pharmaceutical preparation for promoting growth of undifferentiated
CC cells. The medium may also include an inducer, such as insulin, an
CC interleukin, a fibroblast growth factor, or interferon-gamma.
XX
XX Sequence 333 AA:
SQ
Query Match 34.0%; Score 629.5; DB 18; Length 333;
Best Local Similarity 40.9%; Pred. No. 2.1e-43;
Matches 135; Conservative 52; Mismatches 130; Indels 13; Gaps 5;
QY 14 GALLLVGLVGLVSGLEPVMNSANKRFOAGVYLPQIGDRDLDCPRAPPGPHSS 73
DB 14 GVLVLCRTAISKSYLIERIYNSSNSKFLPGQGLVLPQIGDKLDICPKV---DSKTV 70
QY 74 PNYEPKYLVGAGACGRCEAPRNLLTCRPLDLRFTTKFOEYSPNLMGHEFRSH 133
DB 71 GGEYKYVMVWDQADRCTIKKENTPLNCAKPPQDIKFTIKFOEFSBNLMGLEFOQNK 130
QY 134 DVIITATSDGTREGLSELDGGVCLTRGMKVLRLVQ--SPRGAVPRKPVSEMPYER-DR 190
DB 131 DVIITSTNGSLLEGDNOGCVQCTRAKMLMKVQODASASGSTRNKDPTRRPELEAGTN 190
QY 191 GAHSLPEKENVLPDDPTSNATSRGAEGLPPLPSMPAVAGAAGLALLLGVAGAGAMC 250
DB 191 GRSSTSPFVKKNPGSGTGNAGSANNILGSEVALFAGISGCIIFIVIIITLVLL 250
QY 251 WRRRRKAPESHPGSGFRGSGSLGGGGGMGRBAEPGELGALGGGAADPPFCPH 310
DB 251 KTRRRRKSPQHTTTLTSLATATPKSGSN---NGSEPSDIIIPLR---TADSVFCPH 303
QY 311 YEKVSGDYGHPIYIVODGPPQSPNNIYKYV 340
DB 304 YEKVSGDYGHPIYIVQEMPPQSPANNIYKYV 333

```

RESULT 21

AA06334

ID AA06334 standard; protein; 308 AA.

AA06334;

17-JAN-1997 (first entry)

Ligand #2 for receptor-type tyrosine kinase protein.

```

KW Receptor-type tyrosine kinase; ligand; coomassie staining; PAS staining;
KW human.
XX
XX Homo sapiens.
XX
XX JP08188596-A.
XX
XX 23-JUL-1996.
XX
XX 13-JUN-1995; 95JP-0003677.
XX
XX 09-NOV-1994; 94JP-0275411.
XX
XX 19-OCT-1994; 94JP-0253848.
XX
XX (ASAH ) ASAH KASEI KOGYO KK.
XX
XX WPI; 1996-388601/39.
XX
XX New ligand for receptor type tyrosine kinase - has mol.wt. 22-25
PT kilo(dalton)s and is positive for Coomassie and PAS staining
XX
XX Claim 7; Page 44; 51pp; Japanese.
XX
XX AA06333 and AA06334 represent receptor-type tyrosine kinase receptor
CC binding ligands of the invention. These sequences both have the
CC N-terminal sequence represented by AA06332. The ligands of the
CC invention recognise the fragment of the receptor type kinase receptor
CC represented by AA06330 (see AA06331 for full length sequence). The
CC proteins of the invention have a molecular weight of 23500 (plus or minus
CC 1500) Da, and are positive for Coomassie staining and PAS staining. The
CC protein is a new ligand of receptor-type tyrosine kinases, and can be
CC prepared by standard recombinant techniques.
XX
XX Sequence 308 AA:
SQ
Query Match 33.5%; Score 619.5; DB 17; Length 308;
Best Local Similarity 42.0%; Pred. No. 1.2e-42;
Matches 132; Conservative 48; Mismatches 121; Indels 13; Gaps 5;
QY 30 LEPVYNSANKRFOAGVYLPQIGDRDLDCPRAPPGPHSSPNYEFKYLVGAGAG 89
DB 5 LEPIYNSNSKFLPGQGLVLPQIGDKLDICPKV---DSKTVGGYERYKYVMVWDQDA 61
QY 90 RCEBAPRNLLTCRPLDLRFTTKFOEYSPNLMGHEFRSHDYIITATSDGTREGL 149
DB 62 DRCCTIKKENTPLNCAKPPQDIKFTIKFOEFSBNLMGLEFOKQKDYIITSTNGSL 121
QY 150 SLOGGVCLTRGMKVLRLVQ--SPRGAVPRKPVSEMPYER-DRGAHSLPEKENVLP 206
DB 122 NQEGVQCQTRAKMLMKVQODASASGSTRNKDPTRRPELEAGTNGRSTTSPVKENPGS 181
QY 207 PTSNATSRGAEGLPPLPSMPAVAGAAGLALLLGVAGAGAMCWRERRAKPESHPGE 266
DB 182 STDGNSAGSANNILGSEVALFAGISGCIIFIVIIITLVLLKTRRRRKSPQHTT 241
QY 267 GSFGRGSLGGGGGMGRBAEPGELGALGGGAADPPFCPHYKVSVDYGHPIYIVQ 326
DB 242 LSLSTLATPKSGSN---NGSEPSDIIIPLR---TADSVFCPHYKVSVDYGHPIYIVQ 294
QY 327 DGPPQSPNNIYKYV 340
DB 295 EMPQSPANNIYKYV 308

```

RESULT 22

AA04656

ID AA04656 standard; protein; 308 AA.

AA04656;

18-OCT-1996 (first entry)

Ligand for receptor type tyrosine kinase (TK) (mature protein).

```

XX  Receptor type tyrosine kinase; TK; ligand; differentiation;
KM  haematopoietic stem cell; tyrosine; bone marrow; leukaemia.
XX
XX  Homo sapiens.
XX
XX  NO9611212-A1.
XX
XX  18-APR-1996.
XX
XX  09-OCT-1995; 95MO-JP02069.
XX
XX  22-DEC-1994; 94UP-0320712.
PR  07-OCT-1994; 94UP-0244433.
PR  26-OCT-1994; 94UP-0262882.
XX
XX  (ASAH ) ASAH KASEI KOGYO KK.
XX
XX  Ohno M, Sakano S;
XX
XX  WPI, 1996-209809/21.
DR  N-PSDB; AAT18395.
XX
XX  Ligand peptide binding to receptor-type tyrosine kinase - enhances
PT  intracellular tyrosine phosphorylation, useful for investigation of
XX  undifferentiated blood cell behaviour
XX
XX  Disclosure; Page 162-163; 193pp; Japanese.
XX
XX  A ligand polypeptide which binds to the extracellular part of a
CC  specific receptor-type tyrosine kinase and induces phosphorylation
CC  of tyrosine within the cell can be used in the study of the
CC  differentiation of blood cells such as the haematopoietic stem
CC  cells; of disease processes such as leukaemia, and of the biology of
CC  bone marrow transplantation. The ligand plays a role in the
CC  differentiation process and the specific ligand target is expressed
CC  in undifferentiated blood cells.
XX
XX  Sequence 308 AA;
XX
XX  Query Match 33.5%; Score 619.5; DB 17; Length 308;
XX  Best Local Similarity 42.0%; Pred. No. 1.2e-42;
XX  Matches 132; Conservative 48; Mismatches 121; Indels 13; Gaps 5;
XX
XX  30 LEPYVNSAKRFGAEGGYLYPOIGDRDLTLCPRARPPPHSHSPYEPKYLYVGAGG 89
XX  5 LEPYVNSNSKTLPGQGLVLYPQIGDKDITCPKV--DSKTVGQYEVYKVMVDKQA 61
XX
XX  90 RCEAPPAPNLLITCDRFDLRLFTTKFOEYSPNLMGHEFRSHHDYIATSDGTREGLE 149
XX  62 DRCITKXENTPLINCAKPDQDITKTFKFOEYSPNLMGHEFRSHHDYIATSDGTREGLE 121
XX
XX  150 SLOGGVLTGGMKYLAVGQ--SPRGAVPRKRVSEMPMR-DRGAHSLSPCKENLPED 206
XX  122 NQGGVGVQITAMKILMKVGDASAGSTRNKDPTRRPELEAGTNGSSTTSPTVKENPQS 181
XX
XX  207 PTNATSRGAEGPLPPSPMPAVAGAAGLALLLVAGAGAGACWPRRPAKSPSESHPGP 266
XX  182 STDGNSHGSGNNILGSEVALPFIAGISGCTIPIYIITLVLLTKTRRRKHSPPHTTT 241
XX
XX  267 GSGRGGSLDLGGGGGNGPREAPBGLIALRGGAADPPFCHEYKVGSDYGHPIYIQ 326
XX  242 LSLSTLTLTPRSGNN---NGSEPSDIIIPLR--TADSVFCHYKVGSDYGHPIYIQ 294
XX
XX  327 DGPPQSPENIYKY 340
XX  295 EMPQSPANIYKY 308
XX
XX  RESULT 23
XX  AAY71438
XX  AAY71438 standard; peptide; 89 AA.

```

```

AC  AAY71438;
XX
XX  04-OCT-2000 (first entry)
XX
XX  Human ephrin B3 C-terminal cytoplasmic domain.
XX
XX  PDZ domain; B class ephrin; Eph receptor tyrosine kinase; RTK; modulator;
XX  cellular process; nerve cell interaction; regeneration of nerve cell;
XX  axonogenesis; antiproliferative; proliferative disorder; treatment;
XX  differentiative disorder; human; cytoplasmic domain.
XX
XX  Homo sapiens.
XX
XX  Key Location/Qualifiers
XX  Region 57..75
XX  /note= "Conserved region"
XX
XX  Modified-site 60
XX  /note= "Conserved tyrosine phosphorylation site"
XX
XX  Modified-site 67
XX  /note= "Conserved tyrosine phosphorylation site"
XX
XX  Modified-site 72
XX  /note= "Conserved tyrosine phosphorylation site"
XX
XX  Region 78..82
XX  /note= "Conserved region"
XX
XX  Region 84..89
XX  /note= "Conserved region"
XX
XX  Domain 86..89
XX  /label= PDZ_domain
XX
XX  Modified-site 86
XX  /note= "Conserved tyrosine phosphorylation site"
XX
XX  Modified-site 87
XX  /note= "Conserved tyrosine phosphorylation site"
XX
XX  WO200031124-A2.
XX
XX  02-JUN-2000.
XX
XX  19-NOV-1999; 99MO-CA01101.
XX
XX  20-NOV-1998; 98US-0109158.
XX
XX  (MOUN ) MOUNT SINAI HOSPITAL.
XX
XX  Lin D, Pawson A;
XX
XX  WPI; 2000-400038/34.
XX
XX  The patent discloses a complex comprising of a B class ephrin and PDZ
XX  domain containing protein. B class ephrine function as ligands for Eph
XX  receptor tyrosine kinases (RTK) and possesses a transmembrane element and a
XX  highly conserved cytoplasmic tail at the C-terminus, that contains a PDZ
XX  binding site. This complex is used in methods to modulate the interaction
XX  of a B class ephrin and PDZ domain containing protein and to identify
XX  modulator compounds. It is also used for modulating cellular processes
XX  like, axonogenesis, nerve cell interactions and regeneration of nerve
XX  cells. The complex is also useful for treating proliferative or
XX  differentiative disorders associated with this protein complex.
XX  The present sequence is the human ephrin B3 C-terminal cytoplasmic
XX  domain, comprising conserved sequences reminiscent of known or predicted
XX  binding sites for PDZ domains. Ephrin B3 is also known as NLERK-2,
XX  ELK-13, EFL-6, ELF-3 and LERK-8.
XX
XX  Sequence 89 AA;
XX
XX  Query Match 26.6%; Score 492; DB 21; Length 89;
XX  Best Local Similarity 97.8%; Pred. No. 7.5e-33;
XX  Matches 87; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 252 RRRRAKPSRHPGSGFSGSLGLGGGGMGPRAEPGELGIALRGGAADPPCPHY 311
DB 1 RRRRAKPSRHPGSGFSGSLGLGGGGMGPRAEPGELGIALRGGAADPPCPHY 60
QY 312 EKVSGDYGHPVITVODGPPQSPNNIYKY 340
DB 61 EKVSGDYGHPVITVODGPPQSPNNIYKY 89

RESULT 24
ID AAY96782 standard; Protein; 658 AA.
AC AAY96782;
DT 26-SEP-2000 (first entry)
DE Ephrin-B2-Ephrin-B2-FC fusion protein.
KM Ephrin-B2; ELK receptor; ligand; dimer; Fc domain; fusion protein;
XX Efl-6 antagonist; neurological.
OS Chimeric - Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..30
FT /label= Signal peptide
FT /note= "derived from ephrin-B2 ectodomain 1"
FT Protein 31..225
FT /label= Ephrin-B2 Ectodomain-1
FT Peptide 226..228
FT /label= Bridging_peptide
FT Protein 229..423
FT /label= Ephrin-B2 ectodomain_2
FT Peptide 424..426
FT /label= Bridging_peptide
FT Region 427..656
FT /label= Human_IgG1_Fc_region

WO200037642-A1.
29-JUN-2000.
XX 23-DEC-1999; 99WO-US30900.
XX 23-DEC-1999; 99WO-US30900.
XX 23-DEC-1998; 96US-0113387.
XX (REGE-) REGENERON PHARM INC.
XX Davis SJ, Gale NW, Yancopoulos GD, Stahl N;
XX WPI; 2000-442670/38.
XX N-PsDB; AAA51346.
XX Polymulectide encoding a fusion polypeptide, useful for promoting
XX differential function and influencing phenotype, comprises two subunits
XX containing at least one copy of the receptor binding domain of a ligand
XX Example 12; Fig 15A-E; 97pp; English.
XX Production of homogenous forms of clustered ligands is broadly applicable
XX to improve the affinity and/or increase the activity of a ligand as
XX compared to the native form of the ligand. Ephrin fusion proteins have
XX been constructed, which may be useful for treating neurological
XX disorders. The ephrin fusion proteins are preferably capable of binding
XX to ELK receptor and are especially Efl-6 antagonists. The fusion proteins
XX were constructed after it was demonstrated that similar improved
XX activities could be achieved using Tie-2 receptor ligands.
XX Angiopoietin-1 (Ang-1) is a Tie-2 receptor ligand and is an agonist for
XX Tie-2, whereas angiopoietin-2 (Ang-2) is a naturally occurring antagonist
XX of the Tie-2 receptor. The fibrinogen domains (FD) of Ang-1 and Ang-2 are
XX the receptor-binding domains and dimerized versions, of e.g. Ang-1-FD-Fc

```

```

CC (Ang-1 fibrinogen domain fused to an Fc domain), can bind to the Tie-2
CC receptor with much higher affinity than monomeric Ang-1-FD (dimerization
CC occurs between the Fc components of adjacent molecules). However, the Tie-2
CC Ang-1-FD-Fc is not able to induce phosphorylation (activate) the Tie-2
CC receptor on endothelial cells unless it is further clustered with goat
CC anti-human Fc antibodies. The novel fusion proteins, mutant versions of
CC Ang-1-FD and Ang-2-FD, were designed that were intrinsically more highly
CC clustered. Tie-2 agonist fusion proteins may be used as haematopoietic
CC factors. Tie-2 receptor antagonist fusion proteins may be used to
CC diagnose or treat, e.g. myeloproliferative or other proliferative
CC disorders of blood forming organs, e.g. thrombocythemia, polycythemia
CC and leukemias.
XX
XX SQ Sequence 658 AA;
XX
XX Query Match 24.5%; Score 453.5; DB 21; Length 658;
XX Best Local Similarity 34.6%; Pred. No. 1.2e-28;
XX Matches 111; Conservative 32; Mismatches 79; Indels 99; Gaps 8;
QY 5 HSGPGGVRYGALLLGLVLVSGLSLEPYMNSANKRFGAGVLYPQIGRDLDCPR 64
DB 218 HSGNN-----LLGPGIV-----LEPYMNSNKKFPGQGLVLYPQIGDLDCPR 265
QY 65 ARPPGHSSRNYPFTLYLVGAQGRCEAPAPNILLTCDDPDDLRFTTFQGYSPYL 124
DB 266 V--DSKTVGQYRYVYVVDKQADRCTIKENTPLNCAAPDDVKTTFQGRSPYL 322
QY 125 WGEFRSHHDYLIATSDTREGLESIGGVCVLTGKMLTLVVGSPFGCAVRRVSEM 184
DB 323 WGLEPQKNQDYLIISTNSLSLELDQEGVQGTAMKTLKMGVGD----- 368
QY 185 EMERDGAASHLEPKENIPGPTSNATSRGABGPPPSMPAVVGAAGCALLLLGVNG 244
DB 369 -----ASSGASARNHGPFRPELE-----AG 389
QY 245 AGGACMRBRRAKPSRHPGSGFSGSLGLGGGGMGPRAEPGELGIALRGGAAD 304
DB 390 TNG-----RSSITSPFYKPNPQSSITDGNASHSGNNLL-----GGPGE 427
QY 305 P-----PFCPHYKVSQ 316
DB 428 FKSCDHTHTCPCPAPDELGG 448

RESULT 25
ID AAY96781 standard; Protein; 683 AA.
AC AAY96781;
DT 26-SEP-2000 (first entry)
DE Ephrin-B1-Ephrin-B1-FC fusion protein.
XX Ephrin-B1; ELK receptor; ligand; dimer; Fc domain; fusion protein;
XX Efl-6 antagonist; neurological.
XX Chimeric - Homo sapiens.
XX Chimeric - Synthetic.
XX Key Location/Qualifiers
XX Peptide 1..29
XX /label= Signal peptide
XX /note= "derived from ELK-L Ectodomain 1"
XX Protein 30..237
XX /label= ELK-L ectodomain_1
XX Peptide 238..240
XX /label= Bridging_peptide
XX Protein 241..448
XX /label= ELK-L Ectodomain_2
XX Peptide 449..451
XX /label= Bridging_peptide
XX Region 452..683

```

FT XX /label= Human_IgG1_Fc_region
 XX MO200037642-A1.
 XX 29-JUN-2000.
 XX 23-DEC-1999; 99WO-US0900.
 XX 23-DEC-1998; 98US-0113387.
 XX (REGG-) REGENERON PHARM INC.
 XX Davis SJ, Gale NW, Yancopoulos GD, Stahl N;
 XX WPI: 2000-442670/36.
 XX N-PSDB; AAA51345.
 XX Polynucleotide encoding a fusion polypeptide, useful for promoting
 PT differential function and influencing phenotype, comprises two subunits
 PT containing at least one copy of the receptor binding domain of a ligand
 XX
 XX Example 12; Fig 14A-E; 97JP; English.
 XX
 CC Production of homogenous forms of clustered ligands is broadly applicable
 CC to improve the affinity and/or increase the activity of a ligand as
 CC compared to the native form of the ligand. Ephrin fusion proteins have
 CC been constructed, which may be useful for treating neurological
 CC disorders. The ephrin fusion proteins are preferably capable of binding
 CC to Etk receptor and are especially Etk-6 antagonists. The fusion proteins
 CC were constructed after it was demonstrated that similar ligands.
 CC Angiopoietin-1 (Ang-1) is a Tie-2 receptor ligand and is an agonist for
 CC of the Tie-2 receptor. The fibrinogen domains (FD) of Ang-1 and Ang-2 are
 CC the receptor-binding domains and dimerized versions, of e.g. Ang-1-FD-Fc
 CC (Ang-1 fibrinogen domain fused to an Fc domain), can bind to the Tie-2
 CC receptor with much higher affinity than monomeric Ang-1-FD (dimerization
 CC occurs between the Fc components of adjacent molecules). However,
 CC Ang-1-FD-Fc is not able to induce phosphorylation (activated) the Tie-2
 CC receptor on endothelial cells unless it is further clustered with goat
 CC anti-human Fc antibodies. The novel fusion proteins, mutant versions of
 CC Ang-1-FD and Ang-2-FD, were designed that were intrinsically more highly
 CC clustered. Tie-2 agonist fusion proteins may be used as haematopoietic
 CC factors. Tie-2 receptor antagonist fusion proteins may be used to
 CC diagnose or treat, e.g. myeloproliferative or other proliferative
 CC disorders of blood forming organs, e.g. thrombocythemias, polycythemias
 CC and leukemias.
 CC
 XX Sequence 683 AA;
 SQ
 Query Match 24.5%; Score 452.5; DB 21; Length 683;
 Best Local Similarity 44.0%; Pred. No. 1.5e-28;
 Matches 92; Conservative 32; Mismatches 72; Indels 13; Gaps 3;

ID AA024019 standard; Protein; 229 AA.
 XX
 XX AA024019;
 XX 23-SEP-2002 (first entry)
 XX
 XX Murine ephrin B2 ligand.
 XX
 XX Murine; extracellular region; Eph B receptor; ephrin B ligand; tumour;
 XX ocular neovascularisation; gene therapy; cytostatic.
 XX
 XX Mus musculus.
 XX
 XX WO200226627-A1.
 XX
 XX 04-APR-2002.
 XX
 XX 28-SEP-2001; 2001WO-BP11252.
 XX
 XX 29-SEP-2000; 2000CH-0001910.
 XX
 XX (NOVS) NOVARTIS AG.
 XX (NOVS) NOVARTIS-ERRINDUNGEN VERM GES MBH.
 XX Martiny-Baron G, Wood JM, Llau G;
 XX WPI: 2002-405047/43.
 XX N-PSDB; AAD38768.
 XX
 XX New soluble polypeptides of the extracellular region of the Eph B
 XX receptor and ephrin B ligands, useful in gene therapy for treating a
 XX tumor disease or ocular neovascularization in mammals, including humans
 XX
 XX Claim 6; Page 48-49; 55pp; English.
 XX
 XX The invention relates to new isolated soluble polypeptides comprising an
 XX amino acid sequence of an extracellular region of an Eph B receptor or an
 XX amino acid sequence of an ephrin B ligand, or essentially similar amino
 XX acid sequences of extracellular region of an Eph B receptor or an amino
 XX acid sequence of an ephrin B ligand. Extracellular region of an Eph B
 XX receptor is capable of binding with high affinity to a mammalian ephrin B
 XX ligand. Ephrin B ligand is capable of binding with high affinity to a
 XX mammalian Eph B receptor. The nucleic acid encoding the polypeptides or
 XX the polypeptides are useful for producing pharmaceutical compositions for
 XX treating a tumor disease or ocular neovascularisation in mammals,
 XX including humans. The vector is useful in producing a pharmaceutical
 XX composition for treating the diseases cited above in mammals,
 XX particularly humans, by means of gene therapy. The present sequence is
 XX murine ephrin B2 ligand.
 XX
 XX Sequence 229 AA;
 SQ
 Query Match 24.2%; Score 447; DB 23; Length 229;
 Best Local Similarity 36.4%; Pred. No. 1.1e-28;
 Matches 100; Conservative 33; Mismatches 80; Indels 62; Gaps 4;

QY 254 RRAKPSERHPPGSGFRGSGSLGGGGGMGPREA 288
 DB 195 RSSSTSPFVKPNPSSSTDGNSAGHSNNILGSEVA 229

RESULT 27
 ID AAE24020 standard; Protein; 229 AA.
 AC AAE24020;

DT 23-SEP-2002 (first entry)

DE Human ephrin B2 ligand.

KW Human; extracellular region; Eph B receptor; ephrin B ligand; tumour;
 ocular neovascularisation; gene therapy; cyostatic.

OS Homo sapiens.

PN WO200226827-A1.

PD 04-APR-2002.

PF 28-SEP-2001; 2001WO-EP11252.

PR 29-SEP-2000; 2000CH-0001910.

PA (NOVS) NOVARTIS AG.
 (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.

PI Martiny-baron G, Wood JM, Liau G;

DR WPI; 2002-405047/43.
 N-PDSB; AAD38769.

PT New soluble polypeptides of the extracellular region of the Eph B
 receptor and ephrin B ligands, useful in gene therapy for treating a
 tumor disease or ocular neovascularization in mammals, including humans

PS Claim 6; Page 51-52; 55pp; English.

XX The invention relates to new isolated soluble polypeptides comprising an
 CC amino acid sequence of an extracellular region of an Eph B receptor or an
 CC amino acid sequence of an ephrin B ligand, or essentially similar amino
 CC acid sequences of extracellular region of an Eph B receptor or an amino
 CC acid sequence of an ephrin B ligand. Extracellular region of an Eph B
 CC receptor is capable of binding with high affinity to a mammalian ephrin B
 CC ligand. Ephrin B ligand is capable of binding with high affinity to a
 CC mammalian Eph B receptor. The nucleic acid encoding the polypeptides or
 CC the polypeptides are useful for producing pharmaceutical compositions for
 CC treating a tumor disease or ocular neovascularisation in mammals,
 CC including humans. The vector is useful in producing a pharmaceutical
 CC composition for treating the diseases cited above in mammals,
 CC particularly humans, by means of gene therapy. The present sequence is
 CC human ephrin B2 ligand.

SQ Sequence 229 AA;

Query Match 23.9%; Score 443; DB 23; Length 229;

Best Local Similarity 35.6%; Pred. No. 2.4e-28;

Matches 96; Conservative 36; Mismatches 79; Indels 62; Gaps 4;

QY 14 GALLLLGLVGLVSGSLSEPYVNSANRFOAEGGYLYPQIGDRDLDCPRARPSPHSS 73

DB 14 GVLAVLCRTAISKISIVLEPIVYNSNKGFLPGQGLVLYPQIGDKLIIICPKV--DSKTV 70

QY 74 PNYEFLYLYVGAGRGRCCEAPNPNLLTCBPDLDLRTIKFOEYSPYLKMGHFRSHH 133

DB 71 GQYVYVYVWVDQADRCRTIKKENTPLNCAKPPDDIKTIKFOEYSPYLKMGHFRSHH 130

QY 134 DYLIATSDGTREGLELQGGVCLTRGMKVLLVVGOSPRGAVPRKPVSEMPKERVDRGAA 193

DB 131 DYLIATSDGTREGLELQGGVCLTRGMKVLLVVGOSPRGAVPRKPVSEMPKERVDRGAA 166
 QY 194 HSLEPGKENLPDPTSNATSRGAEPLPPSPMPVAGAGLALLILVAGAGMCMRR 253
 DB 167 -----DASAGSTRNKDPTRPELE-----AGTNG----- 191

QY 254 RRAKPSERHPPGSGFRGSGSLGGGGGMGPREA 288
 DB 192 RSSSTSPFVKPNPSSSTDGNSAGHSNNILGSEVA 226

RESULT 28
 ID AAM06333 standard; protein; 195 AA.
 AC AAM06333;

DT 17-JAN-1997 (first entry)

DE Ligand #1 for receptor-type tyrosine kinase protein.

KW Receptor-type tyrosine kinase; ligand; coomassie staining; PAS staining;
 human.

OS Homo sapiens.

PN JP08188596-A.

PD 23-JUL-1996.

PF 13-JAN-1995; 95JP-0003677.

PR 09-NOV-1994; 94JP-0275411.

PR 19-OCT-1994; 94JP-0253848.

PA (ASAH) ASAH KASEI KOGYO KK.

DR WPI; 1996-388601/39.

PT New ligand for receptor type tyrosine kinase - has mol.wt. 22-25
 kilodalton(s) and is positive for Coomassie and PAS staining

PS Claim 5; Page 43; 51pp; Japanese.

XX AAM06333 and AAM06334 represent receptor-type tyrosine kinase receptor
 CC binding ligands of the invention. These sequences both have the
 CC N-terminal sequence represented by AAM06332. The ligands of the
 CC invention recognise the fragment of the receptor type kinase receptor
 CC represented by AAM06330 (see AAM06331 for full length sequence). The
 CC proteins of the invention have a molecular weight of 23500 (plus or minus
 CC 1500) Da, and are positive for Coomassie staining and PAS staining. The
 CC protein is a new ligand of receptor-type tyrosine kinases, and can be
 CC prepared by standard recombinant techniques.

SQ Sequence 195 AA;

Query Match 23.3%; Score 431.5; DB 17; Length 195;

Best Local Similarity 45.3%; Pred. No. 1.7e-27;

Matches 87; Conservative 32; Mismatches 66; Indels 7; Gaps 4;

QY 30 LEPYVNSANRFOAEGGYLYPQIGDRDLDCPRARPSPHSSPNYFXYLYVGAGG 89

DB 5 LEPIYNSNKGFLPGQGLVLYPQIGDKLIIICPKV--DSKTVGYEYKVVWVDQDA 61

QY 90 RRCBAPNPNLLTCBPDLDLRTIKFOEYSPYLKMGHFRSHHDIYIATSDGTREGLE 149

DB 62 DRCITKENTPLNCAKPPDDIKTIKFOEYSPYLKMGHFRSHHDIYIATSDGTREGLE 121

QY 150 SLOGVCLTRGMKVLLVVGQ--SPRGAVPRKPVSEMPKERVDRGAAHSLEPGKENLPD 206

DB 122 NQGGVCLTRGMKVLLVVGQDASAGSTRNKDPTRPELEAGTNGRSTSPFVKPNPSS 181

QY 207 PTSNATSGAAG 218
 DB 182 ST-DGNSAGHSG 192

RESULT 29

AAW11307
 ID AAW11307 standard; protein; 195 AA.

AC AAW11307;

XX 21-MAR-1997 (first entry)

DE Receptor-type tyrosine kinase #3.

XX Receptor-type tyrosine kinase; LPM medium; ligand; autophosphorylation;
 KW insect cell; animal cell; growth promoter; undifferentiated cell;
 KW insulin; interleukin; fibroblast growth factor; hepatocyte growth factor;
 KW nerve growth factor; interferon-gamma; tumour necrosis factor; inducer.

OS Homo sapiens.

PN JF08280385-A.

PD 29-OCT-1996.

PR 13-APR-1995; 95JP-0087878.

PR 13-APR-1995; 95JP-0087878.

PA (ASAH) ASAH KASEI KOGYO KK.

DR WPI; 1997-014848/02.

PT LPM cell culture medium - for culturing cells which produce
 PT receptor-type tyrosine kinase ligand

PS Disclosure; Page 38-39; 45pp; Japanese.

XX AAW11307 and AAW11308 represent receptor-type tyrosine kinases. These
 CC sequences can be included in the LPM medium of the invention. The
 CC medium is used to culture cells capable of producing a ligand for
 CC receptor-type tyrosine kinase. The ligand serves as a physiologically
 CC active substance to induce enzyme activity and autophosphorylation of
 CC tyrosine kinase, which is involved in the development and differentiation
 CC of animal and insect cells. The ligand can also be used as in a
 CC pharmaceutical preparation for promoting growth of undifferentiated
 CC cells. The medium may also include an inducer, such as insulin, an
 CC interleukin, a fibroblast growth factor, or interferon-gamma.

SO Sequence 195 AA;

Query Match 23.3%; Score 431.5; DB 18; Length 195;

Best Local Similarity 45.3%; Pred. No. 1.7e-27;

Matches 87; Conservative 32; Mismatches 66; Indels 7; Gaps 4;

QY 30 LEPTVMSANRFPQAEQYLYPQIGRLDLCPRARPPPHSSPNVEFYKLYLVGAAG 89

DB 5 LEPTVMSNKKFLPGGLVLYPQIGRLDLCPRARPPPHSSPNVEFYKLYLVGAAG 61

QY 90 RRCAPAPNLLTCRPPDLRFTIKFOEYSPNLMGHEFRSHDYIITAGTEGLE 149

DB 62 DRCITKKNPFLNCAKPPDDIKTIKQFSPNLMGHEFRSHDYIITAGTEGLE 121

QY 150 SLQGVCLTRGMKYLAVGQ--SPRGAIVPKVSEPMER-DRGAHSLPCKENLPD 206

DB 122 NQEGVCCOTRAMKILMKVGDASAGSTRNKDPTRRPELAGTNGRSSTSPVKNPFS 181

QY 207 PTSNATSGAAG 218

DB 182 ST-DGNSAGHSG 192

RESULT 30

AAW37671
 ID AAW37671 standard; Protein; 92 AA.

AC AAW37671;

DT 17-OCT-2001 (first entry)

DE Peptide #11708 encoded by probe for measuring placental gene expression.

KW Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.

OS Homo sapiens.

PN W0200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488897/53.

PT Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human placenta -

PS Claim 27; SEQ ID No 37940; 654pp; English.

SO Sequence 92 AA;

Query Match 15.4%; Score 284.5; DB 22; Length 92;

Best Local Similarity 54.2%; Pred. No. 7.4e-16;

Matches 52; Conservative 15; Mismatches 24; Indels 5; Gaps 1;

QY 42 FOAEQYLYPQIGRLDLCPRARPPPHSSPNVEFYKLYLVGAAGRCAPAPNLL 101

DB 1 FLGKGLVLYPQIGRLDLCPRARPPPHSSPNVEFYKLYLVGAAGSTVLDPVVL 55

QY 102 LTCRPPDLRFTIKFOEYSPNLMGHEFRSHDYI 137

DB 56 VTCNRPQEIIRFTIKFOEYSPNLMGHEFRSHDYI 91

RESULT 31

ID ABG46524 standard; Peptide; 92 AA.

AC ABG46524;

DT 19-AUG-2002 (first entry)

DE Human peptide encoded by genome-derived single exon probe SEQ ID 36189.

KW Human; single exon probe; asthma; lung cancer; COPD; ILD;

KM chronic obstructive pulmonary disease; interstitial lung disease;
 KM familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KM Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;
 KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KM primary ciliary dyskinesia; pulmonary hypertension;
 KM hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 XX
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 PS Claim 27; SEQ ID No 36189; 634bp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridization of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridization to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 92 AA;
 SQ
 Query Match 15.4%; Score 284.5; DB 23; Length 92;
 Best Local Similarity 54.2%; Pred. No. 7.4e-16;
 Matches 52; Conservative 15; Mismatches 24; Indels 5; Gaps 1;
 QY 42 FQAEAGVLYTPQIGRDLDCPPAPPPHSSPNEFFYLYLVGAQGRCEAPPAPNLL 101
 DB 1 FLGKGLVTPKIGDKLDLDCPPAENGRP-----YEVYLYLVPRQAAAGCTVLDPNVL 55
 QY 102 LTCDPDLDRFTIKFOEYSPNLMGHEFRSHDPYI 137
 DB 56 VTCNRPQGRFTIKFOEYSPNLMGHEFRSHDPYI 91
 RESULT 32
 ID AAM37534 standard; Protein; 136 AA.
 AC AAM37534;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Peptide #11571 encoded by probe for measuring placental gene expression.
 XX
 KM Probe; microarray; human; placenta; antenatal diagnosis;
 KM genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200157272-A2.
 PN
 PD 09-AUG-2001.
 PF 30-JAN-2001; 2001WO-US00663.
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2001-488897/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 PS Claim 27; SEQ ID No 37803; 654bp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP:
 CC see AAI3315-AA157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 XX
 SQ Sequence 136 AA;
 Query Match 11.1%; Score 205.5; DB 22; Length 136;
 Best Local Similarity 37.7%; Pred. No. 3.5e-09;
 Matches 52; Conservative 13; Mismatches 34; Indels 39; Gaps 5;
 QY 230 GAAGGLA-----LILLGVAGAGA-----MCMRRRRKPSBSRHGP 266

```

Db 11 GASGSSGSDPDGFRNSKALFAAVGAGCVIFLLIIFLTVLLKLRGRKHRTQO----- 65
Qy 267 GSFGRGSLGT-----GGGGGNGPREAREPDELIGIALRGGAADPFPCHYKVSQDYGHVP 322
Db 66 -----RAAALSLSTLASPRGSGGTAGTSPDITIPLR---TTENNYCPHYKVSQDYGHVP 118
Qy 323 YIVODGPPQSPNNIYKYV 340
Db 119 YIVQEMPPQSPNNIYKYV 136

RESULT 33
ABG46394
ID ABG46394 standard; Peptide; 136 AA.
XX
XX ABG46394;
XX
XX 19-AUG-2002 (first entry)
XX
XX Human peptide encoded by genome-derived single exon probe SEQ ID 36059.
XX
XX Human; single exon probe; asthma; lung cancer; COPD, IHD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.
XX
XX Homo sapiens.
XX
XX MO200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US00665.
XX
XX 04-FEB-2000; 2000US-180312P.
XX 26-MAY-2000; 2000US-207456P.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-234687P.
XX 27-SEP-2000; 2000US-236359P.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples -
XX
XX Claim 27; SEQ ID No 36059; 634bp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of
XX probes; the novel set of probes which hybridise at high stringency to a
XX nucleic acid expressed in the human lung; measuring gene expression in a
XX sample derived from human lung, comprising (a) contacting the array with
XX mRNA, and (b) measuring the label detectably bound to each probe of
XX the array; identifying exons in a eukaryotic genome, comprising
XX (a) algorithmically predicting at least one exon from genomic sequences
XX of the eukaryote; and (b) detecting specific hybridisation of detectably

```

```

CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a peptide/protein
CC encoded by a single exon probe of the invention.
CC
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pcl_sequences.
CC
SQ Sequence 136 AA;
CC
CC Query Match 11.1%; Score 205.5; DB 23; Length 136;
CC Best Local Similarity 37.7%; Pred. No. 3.5e-09;
CC Matches 52; Conservative 13; Mismatches 34; Indels 39; Gaps 5;

Qy 230 GAAGGIA-----LLLGAGAGAG-----MCRRRRARRSRRHGP 266
Db 11 GASGSSGSDPDGFRNSKALFAAVGAGCVIFLLIIFLTVLLKLRGRKHRTQO----- 65
Qy 267 GSFGRGSLGT-----GGGGGNGPREAREPDELIGIALRGGAADPFPCHYKVSQDYGHVP 322
Db 66 -----RAAALSLSTLASPRGSGGTAGTSPDITIPLR---TTENNYCPHYKVSQDYGHVP 118
Qy 323 YIVODGPPQSPNNIYKYV 340
Db 119 YIVQEMPPQSPNNIYKYV 136

RESULT 34
AAV71437
ID AAV71437 standard; peptide; 82 AA.
XX
XX AAV71437;
XX
XX 04-OCT-2000 (first entry)
XX
XX Human ephrin B2 C-terminal cytoplasmic domain.
XX
XX PDZ domain; B class ephrin; Eph receptor tyrosine kinase; RTK; modulator;
XX cellular process; nerve cell interaction; regeneration of nerve cell;
XX axonogenesis; antiproliferative; proliferative disorder; treatment;
XX differentiative disorder; human; cytoplasmic domain.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH Region 50..68
XX FT /note= "Conserved region"
XX FT Modified-site 53
XX FT /note= "Conserved tyrosine phosphorylation site"
XX FT Modified-site 60
XX FT /note= "Conserved tyrosine phosphorylation site"
XX FT Modified-site 65
XX FT /note= "Conserved tyrosine phosphorylation site"
XX FT Region 71..75

```

```

FT      /note= "Conserved region"
FT      Region 77..82
FT      /note= "Conserved region"
FT      Domain 79..82
FT      /label= PDZ_domain
FT      Modified-site 79
FT      /note= "Conserved tyrosine phosphorylation site"
FT      Modified-site 80
FT      /note= "Conserved tyrosine phosphorylation site"
XX      WO200031124-A2.
XX      PD 02-JUN-2000.
XX      PF 19-NOV-1999; 99WO-CA01101.
XX      PR 20-NOV-1998; 98US-0109158.
XX      PA (MOUN ) MOUNT SINAI HOSPITAL.
XX      PI Lin D, Pawson A;
XX      DR WPI, 2000-400038/34.
XX      PT Isolated complex for treating proliferative or differentiative
XX      disorders comprises B class ephrin and PDZ domain containing protein -
XX      Example; Fig 1; 59pp; English.
XX      CC The patent discloses a complex comprising of a B class ephrin and PDZ
XX      domain containing protein. B class ephrins function as ligands for Eph
XX      receptor tyrosine kinases (RTK) and possess a transmembrane element and a
XX      highly conserved cytoplasmic tail at the C-terminus, that contains a PDZ
XX      binding site. This complex is used in methods to modulate the interaction
XX      of a B class ephrin and PDZ domain containing protein and to identify
XX      modulator compounds. It is also used for modulating cellular processes
XX      like, axonogenesis, nerve cell interactions and regeneration of nerve
XX      cells. The complex is also useful for treating proliferative or
XX      differentiative disorders associated with this protein complex.
XX      CC The present sequence is the human ephrin B2 C-terminal cytoplasmic
XX      domain, comprising conserved sequences reminiscent of known or predicted
XX      binding sites for PDZ domains. Ephrin B2 is also known as Hck-L, ERF-2,
XX      LERK-5 and NLERK-1.
XX      CC LERK-5 and NLERK-1.
XX      SQ Sequence 82 AA;
XX      Query Match 11.0%; Score 204; DB 21; Length 82;
XX      Best Local Similarity 66.1%; Pred. No. 2.5e-09;
XX      Matches 41; Conservative 4; Mismatches 11; Indels 6; Gaps 2;
OY      279 GGGGMPREAPGELGIALRGGAADPPCPHYEKVSGDYGHPIYIVDGPQSPNNIYK 338
DB      27 GGNNG--SEPSDVIPLR---TADSVCPHYEKVSGDYGHPIYIVDGPQSPANNIYK 80
OY      339 KV 340
DB      81 KV 82
XX      RESULT 35
XX      AAB54187
XX      ID AAB54187 standard; Protein; 106 AA.
XX      AC AAB54187;
XX      DT 09-MAR-2001 (first entry)
XX      DE Human pancreatic cancer antigen protein sequence SEQ ID NO:639.
XX      KW Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
XX      detection; diagnosis; identification; cytoskeletal; neuroprotective;
XX      nontropic; immunomodulatory; relaxant; contractile; gynaecological;
XX      antiinflammatory; cardiac; gene therapy; chromosome mapping;

```

```

KW      linkage analysis; tissue identification; tissue typing; forensic;
KW      neural; immune system; muscular; reproductive; gastrointestinal;
KW      pulmonary; cardiovascular; renal; proliferative.
XX      OS Homo sapiens.
XX      PN WO200055320-A1.
XX      PD 21-SEP-2000.
XX      PF 08-MAR-2000; 2000WO-US05989.
XX      PR 12-MAR-1999; 99US-0124270.
XX      PA (HUMA-) HUMAN GENOME SCI INC.
XX      PI Rosen CA, Ruben SM;
XX      DR WPI; 2000-579444/54.
XX      DR N-PSDB; AAC98952.
XX      PT New nucleic acid that is a pancreatic cancer antigen for preventing,
XX      treating, or ameliorating a medical condition, particular pancreatic
XX      cancer, or for use in assays for diagnosing a pathological condition -
XX      Claim 11; Page 1078; 1379pp; English.
XX      CC AAC98973 to AAC99231 encode the human pancreatic cancer associated
XX      proteins, called pancreatic cancer antigens, given in AAB54008 to
XX      AAB54466. The human pancreatic cancer antigens have cytoskeletal,
XX      neuroprotective, nontropic, immunomodulatory, relaxant, contractive,
XX      gynaecological, cardiac and antiinflammatory activities, and can be used
XX      in gene therapy. The polynucleotide and proteins can be used for
XX      preventing, treating, or ameliorating a medical condition or in assays
XX      for diagnosing a pathological condition or a susceptibility to one in a
XX      subject. Binding partners to the proteins and the activity of the
XX      proteins can be identified. The pancreatic cancer antigens can be used to
XX      detect, treat or prevent pancreatic disorders, especially cancer.
XX      CC Apoptosis and antagonists to the antigens can be screened for. The
XX      pancreatic cancer antigen polynucleotides can be used to design nucleic
XX      acid hybridisation probes that can be used in chromosome mapping, linkage
XX      analysis, tissue identification and/or typing and a variety of forensic
XX      and diagnostic methods. The proteins can be used to generate antibodies
XX      which are used to purify, detect and target the polypeptides, including
XX      CC both in vivo and in vitro diagnostic and therapeutic methods. The
XX      CC proteins can be used to treat or prevent neural, immune system, muscular,
XX      CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
XX      CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
XX      CC sequences used in the exemplification of the present invention.
XX      SQ Sequence 106 AA;
XX      Query Match 10.9%; Score 202.5; DB 21; Length 106;
XX      Best Local Similarity 43.6%; Pred. No. 4.5e-09;
XX      Matches 48; Conservative 11; Mismatches 26; Indels 25; Gaps 4;
OY      235 LALLLGVAAGAGCMCKRRRAKPSBSRHPGSGSLG---GGGGMGPRAAP 220
DB      18 LTVLLRLT-----RKRRKRTQO-----RAALSLSTLAPRGSGGTACTEP 59
OY      291 GELGIALRGGAADPPCPHYEKVSGDYGHPIYIVDGPQSPNNIYK 340
DB      60 SDIIFPLR---TENNCPHYEKVSGDYGHPIYIVDGPQSPANNIYK 106
XX      RESULT 36
XX      AAY71436
XX      ID AAY71436 standard; peptide; 82 AA.
XX      AC AAY71436;
XX      DT 04-OCT-2000 (first entry)
XX      KW

```

DB	Human ephrin B1 C-terminal cytoplasmic domain.
XX	
KW	PDZ domain; B class ephrin; Eph receptor tyrosine kinase; RTK; modulator;
KM	cellular process; nerve cell interaction; regeneration of nerve cell;
RW	axogenesis; antiproliferative; proliferative disorder; treatment;
KV	differential disorder; human; cytoplasmic domain.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Region 50..68 /note= "Conserved region"
FT	Modified-site 53 /note= "Conserved tyrosine phosphorylation site"
FT	Modified-site 60 /note= "Conserved tyrosine phosphorylation site"
FT	Modified-site 65 /note= "Conserved tyrosine phosphorylation site"
FT	Region 71..75 /note= "Conserved tyrosine phosphorylation site"
FT	Region 77..82 /note= "Conserved region"
FT	Region 79..82 /note= "Conserved region"
FT	Domain /label= PDZ_domain
FT	Modified-site 79 /note= "Conserved tyrosine phosphorylation site"
FT	Modified-site 80 /note= "Conserved tyrosine phosphorylation site"
XX	
PV	MO20031124-A2.
XX	
PD	02-JUN-2000.
XX	
PF	19-NOV-1999; 99WO-CA01101.
XX	
PR	20-NOV-1998; 98US-0109158.
XX	
PA	(MOUN) MOUNT SINAI HOSPITAL.
PI	Lin D, Pawson A,
DR	WPI; 2000-400038/34.
XX	
PT	Isolated complex for treating proliferative or differentiative disorders comprises B class ephrin and PDZ domain containing protein -
XX	
PS	Example; Fig 1; 59pp; English.
XX	
CC	The patent discloses a complex comprising of a B class ephrin and PDZ
CC	domain containing protein. B class ephrins function as ligands for Eph
CC	receptor tyrosine kinases (RTK) and possess a transmembrane element and a
CC	highly conserved cytoplasmic tail at the C-terminus, that contains a PDZ
CC	binding site. This complex is used in methods to modulate the interaction
CC	of a B class ephrin and PDZ domain containing protein and to identify
CC	modulator compounds. It is also used for modulating cellular processes
CC	like, axonogenesis, nerve cell interactions and regeneration of nerve
CC	cells. The complex is also useful for treating proliferative or
CC	differential disorders associated with this protein complex.
CC	The present sequence is the human ephrin B1 C-terminal cytoplasmic
CC	domain, comprising conserved sequences reminiscent of known or predicted
CC	binding sites for PDZ domains. Ephrin B1 is also known as LERK-2, Elk-L,
CC	BFL-3, Cer-L and STRAL.
XX	
SQ	Sequence 82 AA;
Query Match	10.9%; Score 201; DB 21; Length 82;
Best Local Similarity	47.3%; Pred. No. 4,4e-09;
Matches 44; Conservative	9; Mismatches 24; Indels 16; Gaps 3
OY	252 RRRRAKPSRRHPGSGSLG-----GGCGNGPREAPRGELGLRGGGAADPPF 307
DB	2 RRRHRHTQQ-----RAAALSTIASPKGSGGTGTESDIILPLR--TTENNY 49

QY	308	CPHEKNSGDYGHPIYTIQDGPQSPNITYYK	340
DB	50	CPHEKNSGDYGHPIYTIQEMPQSPNITYYK	82
RESULT 37			
ABBS8621			
ID	ABBS8621	standard; Protein; 652 AA.	
XX	AC	ABBS8621;	
XX	ABBS8621;		
XX	26-MAR-2002	(first entry)	
XX	Drosophila melanogaster polypeptide SEQ ID NO 2655.		
XX	Drosophila melanogaster polypeptide SEQ ID NO 2655.		
XX	Drosophila; developmental biology; cell signalling; insecticide;		
XX	pharmaceutical.		
XX	Drosophila melanogaster.		
XX	MO200171042-A2.		
XX	27-SEP-2001.		
XX	23-MAR-2001;	2001WO-US09231.	
XX	23-MAR-2000;	2000US-191637P.	
XX	11-JUL-2000;	2000US-0614150.	
XX	(PEKE)	PE CORP NY.	
XX	Venter JC, Adams M, Li FWD, Myers EW,		
XX	WPI; 2001-656860/75.		
XX	N-PSDB; ABL02724.		
XX	New isolated nucleic acid detection reagent for detecting 1000 or more		
XX	genes from Drosophila and for elucidating cell signalling and cell-cell		
XX	interactions -		
XX	Disclousure; SEQ ID NO 2655; 21pp + Sequence listing; English.		
XX	The invention relates to an isolated nucleic acid detection reagent		
XX	capable of detecting 1000 or more genes from Drosophila. The invention is		
XX	useful in developmental biology and in elucidating cell signalling and		
XX	cell-cell interactions in higher eukaryotes for the development of		
XX	insecticides, therapeutics and pharmaceutical drugs. The invention		
XX	discloses genomic DNA sequences (ABL16176-ABL10511), expressed DNA		
XX	sequences (ABBS7737-ABBS72072).		
XX	(ABBS7737-ABBS72072).		
XX	The sequence data for this patent did not form part of the printed		
XX	specification, but was obtained in electronic format directly from WFO		
XX	at ftp.wipo.int/pub/published_pct_sequences.		
XX	Sequence 652 AA;		
QY	Query Match	9.8%; Score 182; DB 22; Length 652;	
DB	Best Local Similarity	22.9%; Pred. No. 1.9e-06;	
DB	Matches 88; Conservative	50; Mismatches 136; Indels 110; Gaps 18;	
QY	33	YVNNNSANRFPQAEQ-GVLYLPQIG-----PRLLDLCPRARPQGHSHSPNTEFYLYLVG	85
DB	219	MHWNNSNIFRIDITDHTDITDKNGLAFEPQVHTIC-VYEPGFENET-EKTIIVNS	276
QY	86	GAQGRRCAPAPA-PNLALTCDRPLDLRFITIKFOYSFNLMGHEFRSHDYYIIATSDGT	144
DB	277	KVEYETGRTIADPRVIAICDKPQKLMFFITTFPBPFPQPGLEPLPNDYYFISTS--S	334
QY	145	REGESLQGGVCLTRGMKIVLAVGQSPRGAVPRKPVSEMEMEDRGAHSL-----EPG	199
DB	335	KDLYRRIRGGRCSITNNKVKVFEKVCACAPEDNN-KTTAUSNKSVTDTGGAINVNTANDES	393

Db 394 HVNHSNNINIAIGTNGGQIIGGPQAGIPINPLSGNNNINGIPTTINSNIDQFNRI 453
 QY 224 SMPAVGAGAGLALLLLGVAGAGAMCWRRAKPSRHPGP-SFGGSLGIGGGG 282
 Db 454 IQPNIIGNHVGNVAVGTGIVGGGIIIL-----TPGAHGININMLQPRGGI 499
 QY 283 MGPRAEPG---ELGIALRG-----GGAADPPFCPHYEKVSGDYGH 320
 Db 500 NG---AYPGHHIQTGIRINNVPTQHNYPHSHKGNANSNINGNDH---HHYNK-----H 547
 QY 321 PYTIYVD-----GPPQSPENIY 337
 Db 548 PNEVVKNEELTYNSGATSDGNIF 571

RESULT 38

ABR65548
 ID ABR65548 standard; Protein; 652 AA.

AC ABR65548;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 23436.

KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL09651.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

PS Dielclosure; SEQ ID NO 23436; 21bp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL10511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABR5737-ABR7072).

CC The sequence data for this patent did not form part of the printed specification, but was contained in electronic format directly from WIPRO at ftp.wipro.int/pub/published_pct_sequences.

CC Sequence 652 AA;

Query Match 9.8%; Score 182; DB 22; Length 652;

Best Local Similarity 22.9%; Pred No. 1.9e-06;

Matches 88; Conservative 50; Mismatches 136; Indels 110; Gaps 16;

QY 33 VYNSANKRQAGG-GVLYYPOIG-----DRLLDLCPRARPPGPHSSPNEYFKLYLVG 85

Db 219 MHNTNSISFRINTDTHIIDVKNLAFFPDVHICP-VYBERGTENET-EKYYIYNS 276
 QY 86 GAQRRCEAPPA-PNLLTCDRBDLDLFTIKQEVSPNLMGHEFRSHDDYIIATSDGT 144
 Db 277 KVEYTCRIYNADPRVIALCDKPOKLMFTITFRPTPOPGGLEFPQNDYFISTSS-8 334
 QY 145 REGLESIGGVCLTRGMKYLIAVQSGPRGAIVRKVSEMPERDGAHSL-----EPG 199
 Db 335 KDDLRYRIGGRCSTNNKRVFVCCAPEDNN-KTYALSNKSVYTDGALINVIANNDS 393
 QY 200 KENLPQDPTSNATSRGAEG-----PLP-----P 223
 Db 394 HVNHSNNINIAIGTNGGQIIGGPQAGIPINPLSGNNNINGIPTTINSNIDQFNRI 453
 QY 224 SMPAVGAGAGLALLLLGVAGAGAMCWRRAKPSRHPGP-SFGGSLGIGGGG 282
 Db 454 IQPNIIGNHVGNVAVGTGIVGGGIIIL-----TPGAHGININMLQPRGGI 499
 QY 283 MGPRAEPG---ELGIALRG-----GGAADPPFCPHYEKVSGDYGH 320
 Db 500 NG---AYPGHHIQTGIRINNVPTQHNYPHSHKGNANSNINGNDH---HHYNK-----H 547
 QY 321 PYTIYVD-----GPPQSPENIY 337
 Db 548 PNEVVKNEELTYNSGATSDGNIF 571

RESULT 39

AAR82605
 ID AAR82605 standard; Protein; 234 AA.

AC AAR82605;

DT 16-MAY-1996 (first entry)

DE Eph transmembrane tyrosine kinase family ligand, Efl-2.

KW Efl-2; EHK1-L; Eph transmembrane tyrosine kinase family ligand; neuroloigical disorder; identification; diagnosis.

OS Homo sapiens.

XX Homo sapiens.

PI Key Location/Qualifiers

FT Peptide 1..30

FT Region 218..235

FT /note= "carboxy terminal hydrophobic GPI-recognition tail"

FT Misc-difference 158 /note= "residue borders main conserved regions"

XX WO9527060-A2.

PD 12-OCT-1995.

PF 04-APR-1995; 95WO-US04208.

PR 21-OCT-1994; 94US-0327423.

PR 04-APR-1994; 94US-0222075.

PR 12-APR-1994; 94US-0229402.

PR 01-SEP-1994; 94US-0299567.

XX (REGF-) REGENERON PHARM INC.

XX Aldrich TH, Davis S, Gale N, Goldfarb M, Maisompierre PC;

XX Yancopoulos GD;

XX WPI; 1995-356635/46.

DR N-PSDB; AAT03883.

PT ligands which bind Eph family receptors - used in the diagnosis of neuroloigical disorders

PS Disclosure; Fig 2; 58pp; English.

XX Efl-2 (also known as EHK-1L) is an Eph transmembrane tyrosine kinase

CC family ligand. It has homology with Bcl (Efl-1) (see AAR2604). Efl-2

CC ends in a C-terminal hydrophobic sequence that appears to be a

CC recognition sequence allowing it to be GPI-linked and thus lacking in

CC an intracellular domain. Efl-2 is useful for identifying other ligands

CC for EHK-1, -2, -3, Eek and Elk receptors. The ligands are useful in

CC promoting a different function and/or influencing the phenotype,

CC such as growth and/or proliferation, of receptor bearing cells. They

CC may be used in the diagnosis, and treatment of neurological disorders.

XX

SQ Sequence 234 AA;

Query Match 9.8%; Score 181; DB 16; Length 234;

Best Local Similarity 28.4%; Pred. No. 6.7e-07;

Matches 64; Conservative 23; Mismatches 82; Indels 56; Gaps 11;

QY 7 GPGGVVAGALLLGLVGLSGLSLBPYVNSANKRFOAGGYLYPGIGRDLDCP--R 64

DB 24 GPG-----GALG-----NRHAVYVNSNQHLRRE-GYVQVNVNDYLDIYCFHYN 67

QY 65 ARPPGPHSSPNVEFYKLYLVGAQGRCEAPAPNILLTCDRPDL--DLRFTIKFOEYS 121

DB 68 SSGAGPDPGGAGAQYLYVNSRNGYRTCAASQFK-RMECNRPAAHSPIKSEKFORYS 126

QY 122 PNLWGHFRSHHDYIYIATSDGTREGLESLOGGVCLTRGKVLRLVQSPRGGAVERPEV 181

DB 127 AFSLGYEFHAGHEYYIISTPTNHL-----WKCLR--MKVFCVCASTSHSG---EKYV 174

QY 182 SEMP-----MERDGAHSLF-----PKENTLP 204

DB 175 PTFMGPNVKNINLEDFEGENPOVPLEKSIKSTSPRREHLF 219

RESULT 40

AAR71481

ID AAR71481 standard; Protein; 238 AA.

XX

AC AAR71481;

XX

DT 03-OCT-1995 (first entry)

XX

DE Human hek-L protein.

XX

KW ligand; cell surface; tyrosine kinase receptor; tumorigenesis;

KW Immunogen.

XX

OS Homo sapiens.

XX

FT Key Location/Qualifiers

FT Peptide 1..19

FT Protein /note= "signal peptide" 20..238

XX

PN WO9506065-A.

XX

PD 02-MAR-1995.

XX

PF 17-AUG-1994; 94WO-US09282.

XX

XX

PR 20-AUG-1993; 93US-0109745.

PR 30-AUG-1993; 93US-0114426.

PR 03-DEC-1993; 93US-0161132.

PR 09-MAY-1994; 94US-0240124.

XX

PA (IMMV) IMMUNEX CORP.

XX

PI Beckmann MP, Cerretti DP;

XX

DR WPI; 1995-106811/14.

DR N-PSDB; AA085887.

XX

PT New isolated DNA encoding hek-L protein or its fusion products

PT useful as assay reagent or for carrying therapeutic and

PT diagnostic compounds to leukemia cells.

XX

PS Claim 21; Page 36; 45pp; English.

XX

CC The sequence is that of a novel protein designated hek-L, a protein

CC that can bind hek (a cell surface receptor tyrosine kinase). Hek-L

CC is the first known ligand for hek and can be used to study cellular

CC processes regulated by hek (which may be involved in tumorigenesis).

CC It is also an immunogen for antibody production, as a reagent for

CC detecting hek or hek-L in vitro assays, to determine binding of

CC hek proteins, to purify hek proteins, and to carry diagnostic or

CC cytotoxic agents to particular leukemia cells that express the hek

CC antigen. Hek-L also binds the elk tyrosine kinase receptors.

CC See also AAR71482.

XX

SQ Sequence 238 AA;

Query Match 9.7%; Score 179; DB 16; Length 238;

Best Local Similarity 28.4%; Pred. No. 1e-06;

Matches 65; Conservative 24; Mismatches 80; Indels 60; Gaps 12;

QY 7 GPGGVVAGALLLGLVGLSGLSLBPYVNSANKRFOAGGYLYPGIGRDLDCP--R 64

DB 24 GPG-----GALG-----NRHAVYVNSNQHLRRE-GYVQVNVNDYLDIYCFHYN 67

QY 65 ARPPGPHSSP---NVEFYKLYLVGAQGRCEAPAPNILLTCDRPDL--DLRFTIKF 117

DB 68 SSGVFGAGPDPGGAGAQYLYVNSRNGYRTCAASQFK-RMECNRPAAHSPIKSEK 126

QY 118 QEYSPNLWGHFRSHHDYIYIATSDGTREGLESLOGGVCLTRGKVLRLVQSPRGGA 177

DB 127 QRYAFSLGYEFHAGHEYYIISTPTNHL-----WKCLR--MKVFCVCASTSHSG--- 174

QY 178 RKPVSSEMP-----MERDGAHSLF-----PKENTLP 204

DB 175 EKPVFTLPQFTMGPNVKNINLEDFEGENPOVPLEKSIKSTSPRREHLF 223

RESULT 41

AAW02587

ID AAW02587 standard; Protein; 184 AA.

XX

AC AAW02587;

XX

DT 28-NOV-1996 (first entry)

XX

DE Lerk-6 protein.

XX

KW Lerk-6; hek; elk; cell surface receptor; culture; reagent;

KW neuron; disorder; injury; delivery agent; diagnostic; therapeutic;

KW Lerk-7; probe; cytokine.

XX

OS Mus musculus.

XX

PN WO9617925-A1.

XX

PD 13-JUN-1996.

XX

PF 05-DEC-1995; 95WO-US15781.

XX

XX

PR 01-MAR-1995; 95US-0396946.

PR 06-DEC-1994; 94US-0351025.

XX

PA (IMMV) IMMUNEX CORP.

XX

PI Cerretti DP;

XX

DR WPI; 1996-287171/29.

DR N-PSDB; AAT32700.

XX

PT New isolated human Lerk-7 cytokine - which binds to cell surface

PT receptors elk, hek and eck, useful for delivering agents to cells or
for treating neural disorders

Example 1; Page 35; 49pp; English.

CC The present sequence is that of a murine cytokine, Lerk-6, encoded by
CC AAT32700. A fragment of murine Lerk-6 DNA was isolated by PCR for use as
CC a probe for a human Lerk-6 homologue. The probe however, led to the
CC identification of a human Lerk-7 gene (see AAT32699). The Lerk-7 gene
CC and protein can be used for studying the role of Lerk-7 in conjunction
CC with elk, hek and eck receptors. They can also be used for delivering
CC diagnostic or therapeutic agents to cells, e.g. cancer cells. The
CC Lerk-7 proteins can also exhibit neuroprotective or neurotrophic
CC properties and can be used to treat neural tissue disorders.

CC Sequence 184 AA;

Query Match 9.5%; Score 176; DB 17; Length 184;
Best Local Similarity 29.3%; Pred. No. 1.3e-06;
Matches 58; Conservative 19; Mismatches 69; Indels 52; Gaps 7;

QY 33 VYMSANKRFQAE-----GGYLYPQIGDRDLCPAPAPGPHSPNYEFYKLYVGGA 87
DB 10 VYMSNPNRQVAVGDDGGYTVESINDYLDYCPHYGAPLP-PAERMERYILYVWNGE 68
QY 88 QGRRCAPAPAPNLLTCDRPLD---DLRTIKFQEVSPMLMGHEFRSHDYIATSDGT 144
DB 69 GHASCDHRGRGFRKWEKCNPPAPGGLKSEKQFTFPSLGFERPGEHYIISATP-- 126
QY 145 REGLESLOGGVCLTRGMKYLRYGQSPRGAAVPRKVSSEMPERDGAHSLPEKKNLP 204
DB 127 ----PMLVDRPCLR--LKYVVR-----PTNETLY 149
QY 205 GDP-----TSNATSRGAE 218
DB 150 EAPEPIFTSNSSCSGLGG 167

RESULT 42

AAR91283
ID AAR91283 standard; Protein; 184 AA.

AC AAR91283;
DT 10-OCT-1996 (first entry)
XX
DE LERK-6 protein.

KM LERK-6; hek; elk; cell surface receptors; culture; reagent;
KW neurons; disorder; injury; delivery agent; diagnostic; therapeutic.

OS Mus musculus.

PN MO9610911-A1.

PD 18-APR-1996.

PF 04-OCT-1995; 95WO-US12779.

PR 03-OCT-1995; 95US-0538709.

PR 05-OCT-1994; 94US-0318393.

PA (IMMV) IMMUNEX CORP.

PI Cerretti DP;

DR WPI; 1996-209575/21.

DR N-PSDB; AAT14009.

PT Isolated DNA encoding cytokine designated LERK-6 which binds to hek
PT and elk cell surface receptors - useful for drug delivery and
PT screening procedures.

PS Claim 6; Page 34; 44pp; English.

CC The LERK-6 polypeptide encoded can be used to isolate cells
CC expressing hek/elk cell surface receptors, or to measure the
CC biological activity of such receptors. The protein may also be used
CC as a delivery agent, taking diagnostic and therapeutic agents to
CC cells expressing such receptors. LERK-6 can also be used as a tissue
CC culture reagent to enhance the viability or prolong the lifespan of
CC the neurons. Neural tissue disorders and injuries may be treated by
CC contact with the polypeptide.

CC Sequence 184 AA;

Query Match 9.5%; Score 176; DB 17; Length 184;
Best Local Similarity 29.3%; Pred. No. 1.3e-06;
Matches 58; Conservative 19; Mismatches 69; Indels 52; Gaps 7;

QY 33 VYMSANKRFQAE-----GGYLYPQIGDRDLCPAPAPGPHSPNYEFYKLYVGGA 87
DB 10 VYMSNPNRQVAVGDDGGYTVESINDYLDYCPHYGAPLP-PAERMERYILYVWNGE 68
QY 88 QGRRCAPAPAPNLLTCDRPLD---DLRTIKFQEVSPMLMGHEFRSHDYIATSDGT 144
DB 69 GHASCDHRGRGFRKWEKCNPPAPGGLKSEKQFTFPSLGFERPGEHYIISATP-- 126
QY 145 REGLESLOGGVCLTRGMKYLRYGQSPRGAAVPRKVSSEMPERDGAHSLPEKKNLP 204
DB 127 ----PMLVDRPCLR--LKYVVR-----PTNETLY 149
QY 205 GDP-----TSNATSRGAE 218
DB 150 EAPEPIFTSNSSCSGLGG 167

RESULT 43

AAV06820
ID AAV06820 standard; Protein; 184 AA.

AC AAV06820;

DT 24-JUN-1999 (first entry)

XX Murine LERK-6 polypeptide.

KM LERK-6 polypeptide; hek receptor; elk receptor; human; murine; injury;
KW cell proliferation; neural growth; neural tissue; neurological disease;
KM neurodegenerative; excitotoxicity.

OS Mus sp.

PN WO910495-A1.

PD 04-MAR-1999.

PF 27-AUG-1998; 98WO-US17772.

PR 29-AUG-1997; 97US-0920440.

PA (IMMV) IMMUNEX CORP.

PI Cerretti DP;

DR WPI; 1999-243567/20.

DR N-PSDB; AAX32761.

PT New cytokine designated LERK-6

PS Claim 6; Page 37-38; 46pp; English.

CC The invention relates murine and human LERK-6 polypeptides that bind to
CC hek/elk receptors. Host cells transfected or transformed with vectors
CC comprising the LERK-6 nucleic acid sequences are used for the recombinant
CC production of the proteins. LERK-6 polypeptides may be useful in the


```

FT      /note= "potential N-linked glycosylation site"
FT      /note= 35..166
FT      /note= "contains core sequence motif"
FT      /note= 69..159
FT      /note= "contains a Cys4 motif"
FT      /note= 170
FT      /note= "potential N-linked glycosylation site"
FT      /note= 184
FT      /note= "potential N-linked glycosylation site"
XX      US5795734-A.
XX      18-AUG-1998.
XX      31-MAY-1995; 95US-0455001.
XX      31-MAY-1995; 95US-0455001.
XX      19-SEP-1994; 94US-0308814.
XX      27-FEB-1995; 95US-0393462.
XX      (HARD ) HARVARD COLLEGE.
XX      PI      Cheng H, Flanagan JG;
XX      DR      WPI; 1998-46665/40.
XX      DR      N-PSDB; AAY42926.
XX      PT      Nucleic acid encoding E1f-1 protein that binds to EPH-type receptor
XX      PT      - for production of E1f-1 protein, useful for regulating
XX      PT      proliferation, differentiation, and survival of cells
XX      PS      Claim 1; Fig 2A; 53pp; English.
XX      CC      The present sequence represents a mammalian EPH receptor ligand
XX      CC      designated E1f-1. This ligand can bind to both mex-4 and sek-xp. E1f-1 is
XX      CC      a tyrosine kinase ligand, which is linked to the membrane through a
XX      CC      phosphatidylinositol linkage. It shares some homology to 2 other EPH
XX      CC      receptor ligands, B61 and LERK-2. The E1f-1 protein modulates
XX      CC      proliferation, differentiation and survival of EPH receptor-expressing
XX      CC      cells by stimulating or antagonising intracellular signalling mediated
XX      CC      by the EPH receptor. Typical of many potential applications are
XX      CC      increasing survival of neuronal cells in culture (e.g. where intended for
XX      CC      transplantation), also therapeutically in increase neuron survival
XX      CC      (e.g. treatment of Alzheimer's or Parkinson's diseases), to prevent
XX      CC      nervous system and lymphatic tumours, to induce differentiation of
XX      CC      hepatocytes to form an artificial liver, to induce cartilage and bone
XX      CC      formation.
XX      SQ      Sequence 209 AA;
XX      Query Match 9.5%; Score 176; DB 19; Length 209;
XX      Best Local Similarity 29.3%; Pred. No. 1.5e-06;
XX      Matches 58; Conservative 19; Mismatches 69; Indels 52; Gaps 7;
XX      QY      33 VYVNSANKRQAE-----GGVLYVPOIGDRLDLCPRARPSPGHSPPNVEFYLYVGA 87
XX      DB      35 VYVNSRPRQVSAVDGGGTYEVSINDYLDYCPHYGAPLP-PARKERYILLVWNGE 93
XX      QY      88 QGRCEAPPAANLLITCDRDL--DLFTIKQEQYSFNLWGHEFRSHDYIATSDET 144
XX      DB      94 GHASCDHRQGFQKMECNRPAPAPGGLPKFSEKQLTFPFLGFEFRPGHEYYIISATP-- 151
XX      QY      145 REGLESIGGVCLTRGKNTLLRVQSGPRGAVRKPVSSEPMERDGAHSLDEPGENLP 204
XX      DB      152 ----PNLVDRPCLR--LKVYVR-----PMTVELLY 174
XX      QY      205 GDP-----TSNATSRGAG 218
XX      DB      175 EAPPEPFTSNSSCSGLG 192

```

RESULT 46
AAY06822

```

ID      AAY06822 standard; Protein; 213 AA.
XX      AC      AAY06822;
XX      DT      24-JUN-1999 (first entry)
XX      DE      Human LERK-6 polypeptide.
XX      KM      LERK-6 polypeptide; hek receptor; elk receptor; human; murine; injury;
XX      KM      cell proliferation; neural growth; neural tissue; neurological disease;
XX      KM      neurodegenerative; excitotoxicity.
XX      OS      Homo sapiens.
XX      PN      W09J10495-A1.
XX      PD      04-MAR-1999.
XX      PF      27-AUG-1998; 98MO-US17772.
XX      PR      29-AUG-1997; 97US-0920440.
XX      PA      (IMMUNEX ) IMMUNEX CORP.
XX      PI      Corrected DP;
XX      DR      WPI; 1999-243567/20.
XX      DR      N-PSDB; AAX32767.
XX      PT      New cytokine designated LERK-6
XX      PS      Claim 6; Page 42; 46pp; English.
XX      CC      The invention relates murine and human LERK-6 polypeptides that bind to
XX      CC      hek/elk receptors. Host cells transfected or transformed with vectors
XX      CC      comprising the LERK-6 nucleic acid sequences are used for the recombinant
XX      CC      production of the proteins. LERK-6 polypeptides may be useful in the
XX      CC      enhancement, stimulation, proliferation or growth of cells expressing the
XX      CC      hek or elk receptor. The ligand and receptor complex may be involved in
XX      CC      neural growth, development and/or maintenance. LERK-6 can be used for
XX      CC      treating disorders of neural tissue such as injury or neurological
XX      CC      diseases, either chronic or acute. LERK-6 may be employed in treating
XX      CC      neurodegenerative conditions where there is neural death, excitotoxicity.
XX      CC      In addition, they may be administered to a mammal to exert a trophic
XX      CC      effect on neural tissue. They can also be used as reagents for those
XX      CC      conducting quality assurance studies e.g. to monitor shelf life and
XX      CC      stability of elk protein under different conditions. The polypeptides can
XX      CC      also be used as carriers for delivering agents attached to cells bearing
XX      CC      the elk or hek cell surface receptor. The present sequence represents a
XX      CC      human LERK-6 polypeptide.
XX      SQ      Sequence 213 AA;
XX      Query Match 9.5%; Score 175.5; DB 20; Length 213;
XX      Best Local Similarity 36.8%; Pred. No. 1.7e-06;
XX      Matches 43; Conservative 14; Mismatches 51; Indels 9; Gaps 3;
XX      QY      33 VYVNSANKRQAE-----EGVLYVPOIGDRLDLCPRARPSPGHSPPNVEFYLYVGA 87
XX      DB      39 VYVNSRPRQVSAVDGGGTYEVSINDYLDYCPHYGAPLP-PARKERYILLVWNGE 97
XX      QY      88 QGRCEAPPAANLLITCDRDL--DLFTIKQEQYSFNLWGHEFRSHDYIATSDET 141
XX      DB      98 GHASCDHRQGFQKMECNRPAPAPGGLPKFSEKQLTFPFLGFEFRPGHEYYIISAT 154

```

RESULT 47
AAB54058
ID AAB54058 standard; Protein; 218 AA.
XX AAB54058;
XX AC AAB54058;
XX DT 09-MAR-2001 (first entry)

DE	XX	Human pancreatic cancer antigen protein sequence SEQ ID NO:510.
XX	XX	Human, pancreas; pancreatic cancer; pancreatic cancer antigen;
XX	XX	detection; diagnosis; identification; cytostatic; neuroprotective;
XX	XX	nootropic; immunomodulatory; relaxant; contraceptive; gynaecological;
XX	XX	antitumorigenic; cardiac; gene therapy; chromosome mapping;
XX	XX	linkage analysis; tissue identification; tissue typing; forensic;
XX	XX	neural; immune system; muscular; reproductive; gastrointestinal;
XX	XX	pulmonary; cardiovascular; renal; proliferative.
OS	XX	Homo sapiens.
XX	XX	WO20055320-A1.
XX	XX	21-SEP-2000.
XX	XX	08-MAR-2000; 2000NO-US05989.
XX	XX	12-MAR-1999; 99US-0124270.
XX	XX	(HUMA-) HUMAN GENOME SCI INC.
XX	XX	Rosen CA, Ruben SM;
XX	XX	WPI, 2000-579444/54.
XX	XX	N-PSDB; AAC98823.
XX	XX	New nucleic acid that is a pancreatic cancer antigen for preventing,
XX	XX	treating, or ameliorating a medical condition, particular pancreatic
XX	XX	cancer, or for use in assays for diagnosing a pathological condition -
XX	XX	Claim 11; Page 946-947; 1379pp; English.
XX	XX	AAC98773 to AAC99231 encode the human pancreatic cancer associated
XX	XX	proteins, called pancreatic cancer antigens, given in AAB54008 to
XX	XX	AAB54466. The human pancreatic cancer antigens have cytostatic,
XX	XX	neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
XX	XX	gynaecological, cardiant and antitumorigenic activities, and can be used
XX	XX	in gene therapy. The polynucleotide and proteins can be used for
XX	XX	preventing, treating, or ameliorating a medical condition or in assays
XX	XX	for diagnosing a pathological condition or a susceptibility to one in a
XX	XX	subject. Binding partners to the proteins and the activity of the
XX	XX	proteins can be identified. The pancreatic cancer antigen can be used to
XX	XX	detect, treat or prevent pancreatic disorders, especially cancer.
XX	XX	Agonists and antagonists to the antigens can be screened for. The
XX	XX	pancreatic cancer antigen polynucleotides can be used to design nucleic
XX	XX	acid hybridization probes that can be used in chromosome mapping, linkage
XX	XX	analysis, tissue identification and/or typing and a variety of forensic
XX	XX	and diagnostic methods. The proteins can be used to generate antibodies
XX	XX	which are used to purify, detect and target the polypeptides, including
XX	XX	both in vivo and in vitro diagnostic and therapeutic methods. The
XX	XX	proteins can be used to treat or prevent neural, immune system, muscular,
XX	XX	reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
XX	XX	proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
XX	XX	sequences used in the exemplification of the present invention.
XX	XX	Sequence 218 AA;
XX	XX	9.3%; Score 172.5; DB 21; Length 218;
XX	XX	Best Local Similarity 28.3%; Pred. No. 3.1e-06;
XX	XX	Matches 51; Conservative 30; Mismatch 80; Indels 19; Gaps 6
XX	XX	8 PGCVFVGLLLGLVGLVSGSL---EPYVNSANRFGAEGYLVYPIQIDRLDLCPR 64
XX	XX	8 PQQAAMFELWAPLILGICCSLAADRHTVWNSNKPFRNE-DYTHQVLDYDILDPH 66
XX	XX	65 AARPPEHSSPN--YEFYQLYVGAQGRCEAPAPNLLTCDRPDL--DLRFTIXQE 119
XX	XX	67 YE---DHRVADAAMEQVILYVHEHYEQIQPQSKQVYAMQCNRSPAQKHGPEKLSKFOR 123
XX	XX	120 VSPNMGHEFHSMDHYIYIANSQTRGCLSLDSQGVCIIRKMKYLLRVQSGRGANVRK 179

```
Db      124 FTPFTLGFEPFEGHSYYSKPIQHEDR-----CLRKLVTVSQKTTSPQAHNDPQE 176
RESULT 48
AA94767 standard; Protein; 200 AA.
ID     AAR94767
XX     AAR94767;
AC     02-JUL-1996 (first entry)
D       Chicken EPH receptor ligand Elf-1.
DE     Bif-1; Eph receptor ligand; dementia; tachycardia; therapy;
KW     diagnosis; transgenic animal.
XX
OS     Gallus sp.
FH     Key Location/Qualifiers
FT     Peptide 1..21
FT         /label= sig_peptide
FT         /label= Mat_protein
FT         /label= Cys4_motif
FT     Region 61..150
FT         /label= Core_sequence_motif
FT         /label= 33..157
FN     MO9609384-AI.
XX PD    28-MAR-1996.
XX PF    19-SEP-1995; 95WO-US11869.
XX PR    27-FEB-1995; 95US-0393462.
XX PR    19-SEP-1994; 94US-C0308414.
PA      (HARD ) HARVARD COLLEGE.
PI      Cheng H, Flanagan JG;
DR      WPI; 1996-188446/19.
N-PSDB; AAT15009.
PT      Murine and chicken EPH receptor ligand, Elf-1 - useful in diagnosis
PT      and treatment of disorders associated with the Elf-1 gene, e.g.
PT      dementia, tachycardia , etc.
PS      Claim 1; Page 88-89; 107pp; English.
CC      A novel chicken EPH receptor ligand, Elf-1 (AAR94767), is involved in
CC      the formation and maintenance of ordered spatial arrangements of
CC      differentiated tissue. It is the product of a cDNA clone (AAT15009)
CC      obtained from an embryo cDNA expression library. This cDNA is used for
CC      the production of recombinant Elf-1, which can be used to modulate
CC      proliferation, survival and/or differentiation of cells and tissues,
CC      and to stimulate or antagonise intracellular signal transduction
CC      pathways mediated by the EPH-type receptor.
SQ      Sequence 200 AA;
Query Match 93%, Score 172; DB 17; Length 200;
Best Local Similarity 35.3%; Pred.No.3e-06; Mismatches 58; Indels 16; Gaps
Matches 49; Conservative 16;
OY      15 ALLILGLVLGVLSGLSLER-----YYWNSANKRFQAGGGVLYLPQIGRLDLLCPRA 65
           ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DY      7 AALLIAIVG-VCVMSDDDPKVISDRYAAYVMRSNPRFH-RGDYTVEVISINDYLDIYCPTY 64
OY      66 RPPGHSSSNVFYLTLLYGAGRCRAPRNILLTCDRPDL---DLRFTIKFOEYXP 122
           :|||:|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DY      65 EEPLP--RRMERRYLVIMWNTEGHASCDHRQGCFRWECNRDSGSGLUKSEKQLTFP 122
```

QY 123 NLWGHEFRSHDYIATIS 141
 DB 123 FSLGFEFRPGHEYYIAS 141

RESULT 49
 AAW71007
 ID AAW71007 standard; Protein; 200 AA.

AAW71007;

20-OCT-1998 (first entry)

XX Amino acid sequence of an avian E1f-1 protein.

XX Chicken; EPH receptor ligand; E1f-1; mek-4; sek-AF;
 KW tyrosine kinase ligand; B61; LERK-2; proliferation; differentiation;
 KW intracellular signalling; increased; survival; neuronal cell;
 KW neuron survival; treatment; Alzheimer's; Parkinson's; lymphatic tumour;
 KW artificial liver; cartilage; bone formation.

XX Gallus sp.

XX Key Location/Qualifiers

XX Misc-difference 61..150
 FT note= "contains a Cys4 motif"

XX Misc-difference 35..157
 FT note= "contains a core sequence motif"

XX US5795734-A.

XX 18-AUG-1998.

XX 31-MAY-1995; 95US-0455001.

XX 31-MAY-1995; 95US-0455001.

XX 19-SEP-1994; 94US-0308814.

XX 27-FEB-1995; 95US-0393462.

XX (HARD) HARVARD COLLEGE.

XX Cheng H, Flanagan JG;

XX WPI; 1998-46665/40.

XX N-PSDB; AAY42927.

XX Nucleic acid encoding E1f-1 protein that binds to EPH-type receptor

XX - for production of E1f-1 protein, useful for regulating

XX proliferation, differentiation, and survival of cells

XX Claim 1; Columns 75-78; 53pp; English.

XX The present sequence represents an avian EPH receptor ligand designated
 CC E1f-1. This ligand can bind to both mek-4 and sek-AP. E1f-1 is a
 CC tyrosine kinase ligand, which is linked to the membrane through a
 CC phosphatidylinositol linkage. It shares some homology to 2 other EPH
 CC receptor ligands, B61 and LERK-2. The E1f-1 protein modulates
 CC proliferation, differentiation and survival of EPH receptor-expressing
 CC cells by stimulating or antagonising intracellular signalling mediated
 CC by the EPH receptor. Typical of many potential applications are
 CC increasing survival of neuronal cells in culture (e.g. where intended for
 CC transplantation), also therapeutically in increase neuron survival
 CC (e.g. treatment of Alzheimer's or Parkinson's diseases), to prevent
 CC nervous system and lymphatic tumours, to induce differentiation of
 CC hepatocytes to form an artificial liver, to induce cartilage and bone
 CC formation.

XX Sequence 200 AA;

Query Match 9.3%; Score 172; DB 19; Length 200;
 Best Local Similarity 35.3%; Pred. No. 3e-06; 58; Indels 16; Gaps 5;
 Matches 49; Conservative 16; Mismatches 58; Indels 16; Gaps 5;

QY 15 ALLLGLVGLVSGLSLEP-----VYVNSANKRFOAEGGYLYXPOIGRLDILCPRA 65
 DB 7 AALLAIVG-VGVMSDDDEKATSDRIYAVYVNSNRFH-RGDTYVAVINDYLDYCHY 64
 QY 66 RPPGHSSENYEYLYLVGAGRCCEAPAPNLLITCDPDL--DLFTIKQESYSP 122
 DB 65 EEPLP--AERMERVLYLVVYVNEGHASCDHRQGFGRMECNRPDPSGDLKSEKQLEPTP 122

QY 123 NLWGHEFRSHDYIATIS 141
 DB 123 FSLGFEFRPGHEYYIAS 141

RESULT 50

AA71482
 ID AAR71482 standard; Protein; 201 AA.

AA71482;

03-OCT-1995 (first entry)

XX Human hek-L protein.

XX Ligand; cell surface; tyrosine kinase receptor; tumorigenesis;
 KW immunogen.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1..22
 FT note= "signal peptide"

XX Protein 23..201

XX WO9506065-A.

XX 02-MAR-1995.

XX 17-AUG-1994; 94WO-US09282.

XX 20-AUG-1993; 93US-0109745.

XX 30-AUG-1993; 93US-0114426.

XX 03-DEC-1993; 93US-0161132.

XX 09-MAY-1994; 94US-0240124.

XX (IMM) IMMUNEX CORP.

XX Beckmann MP, Cerretti DP;

XX WPI; 1995-106811/14.

XX N-PSDB; AAO85888.

XX New isolated DNA encoding hek-L protein or its fusion products

XX useful as assay reagent or for carrying therapeutic and

XX diagnostic compounds to leukemia cells.

XX Claim 21; Page 38; 45pp; English.

XX The sequence is that of a novel protein designated hek-L, a protein
 CC that can bind hek (a cell surface receptor tyrosine kinase). Hek-L
 CC is the first known ligand for hek and can be used to study cellular
 CC processes regulated by hek (which may be involved in tumorigenesis).
 CC It is also an immunogen for antibody production, as a reagent for
 CC detecting hek or hek-L in vitro assays, to determine binding of
 CC hek proteins, to purify hek proteins, and to carry diagnostic or
 CC cytotoxic agents to particular leukemia cells that express the hek
 CC antigen. Hek-L also binds the elk tyrosine kinase receptors.

XX See also AAR71481.

XX Sequence 201 AA;

Query Match 9.2%; Score 169.5; DB 16; Length 201;
 Best Local Similarity 29.9%; Pred. No. 4.9e-06;
 Matches 66; Conservative 18; Mismatches 82; Indels 55; Gaps 10;

```

QY 24 LVGLSL-EPYVMSANRFPAGEGYLYLPQIGRLDLCRRAPRPPSHSPNYEYKY 82
Db 20 LRGSSLEHHVVVYVMSNRL-LRGDAVELGLNDLYLDVCPHGEPPGSP--ETPALT 76
QY 83 LVGAGGRCRAP-PAPNLLITCDRPDLDTFTIKFOEYSPNLWGHEFRSHHDYIIATS 141
Db 77 MYDMGYESCGAEGPRAKRWVCSLFFGHVQFSKIQRTFFSLGFEPFLPGETYYIATP 136
QY 142 DGTREGLESLGGCVLTLMKMYLRLVGSPRGCAVPKRPVSEMEMEDRGAASLEBCKE 201
Db 137 --TPE-----SSGQL-----RQVSVCKEKESASAPV----- 164
QY 202 NLPGDPTSNATS--RGAEGPLPPEMPAVAGAAGLALLIL 240
Db 165 ---GSPGESGTSGMKGDTPSP-----LCILL 189

```

Search completed: February 11, 2003, 12:03:00
Job time : 32.2264 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 12:01:50 / Search time 10.6918 Seconds
(without alignments)
935.649 Million cell updates/sec

Title: US-10-021-121-4

Perfect score: 1850
Sequence: 1 MGPPHSGGCGVGVGALLLLG.....PYIVODGPPGSPNNYYKV 340

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database: Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfillset1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1846	99.8	340	4	US-09-214-631-3
2	637.5	34.5	336	1	US-08-436-044-2
3	637.5	34.5	336	2	US-08-436-054-2
4	637.5	34.5	336	5	PCT-US95-08812-2
5	632	34.2	346	1	US-08-213-403-2
6	632	34.2	346	1	US-08-458-077-2
7	632	34.2	346	1	US-08-460-741-2
8	632	34.2	346	1	US-08-747-240-2
9	632	34.2	346	1	US-08-299-567-6
10	629.5	34.0	333	1	US-08-436-044-4
11	629.5	34.0	333	2	US-08-436-054-4
12	629.5	34.0	333	4	US-08-271-948-2
13	629.5	34.0	333	5	PCT-US95-08534-2
14	629.5	33.6	333	5	PCT-US95-08812-4
15	622.5	33.6	333	4	US-09-214-631-4
16	608.5	32.9	345	4	US-09-214-631-5
17	498	26.9	89	4	US-09-214-631-13
18	201	10.9	82	4	US-09-214-631-11
19	200.5	10.8	82	4	US-09-214-631-12
20	179	9.7	234	4	US-08-299-567-5
21	179	9.7	238	1	US-08-240-124-2
22	179	9.7	238	1	US-08-453-943-2
23	179	9.7	238	2	US-09-057-121-2
24	179	9.7	238	4	US-09-358-734-2
25	176.5	9.5	135	1	US-08-299-567-7
26	176	9.5	184	1	US-09-609-324A-2
27	176	9.5	184	2	US-08-920-440B-2

28	176	9.5	184	4	US-09-173-492-2	Sequence 2, Appl1
29	176	9.5	184	4	US-09-173-133-2	Sequence 2, Appl1
30	176	9.5	184	4	US-09-165-533-2	Sequence 2, Appl1
31	176	9.5	184	5	PCT-US95-12779-2	Sequence 2, Appl1
32	176	9.5	184	5	PCT-US95-15781-2	Sequence 2, Appl1
33	176	9.5	209	1	US-08-455-001-2	Sequence 2, Appl1
34	176	9.5	209	4	US-08-308-814-2	Sequence 2, Appl1
35	176	9.5	209	5	PCT-US95-11869-2	Sequence 2, Appl1
36	175.5	9.5	213	1	US-09-609-324A-10	Sequence 10, Appl1
37	175.5	9.5	213	2	US-08-920-440B-10	Sequence 10, Appl1
38	175.5	9.5	213	4	US-09-173-492-10	Sequence 10, Appl1
39	175.5	9.5	213	4	US-09-173-133-10	Sequence 10, Appl1
40	174.5	9.4	233	4	US-09-214-631-7	Sequence 7, Appl1
41	172	9.3	200	5	US-08-455-001-4	Sequence 4, Appl1
42	172	9.2	200	5	PCT-US95-11869-4	Sequence 4, Appl1
43	169.5	9.2	201	1	US-08-240-124-4	Sequence 4, Appl1
44	169.5	9.2	201	1	US-08-453-943-4	Sequence 4, Appl1
45	169.5	9.2	201	2	US-09-057-121-4	Sequence 4, Appl1
46	169.5	9.2	201	4	US-09-358-734-4	Sequence 4, Appl1
47	169	9.1	201	4	US-09-214-631-8	Sequence 8, Appl1
48	168.5	9.1	209	4	US-09-214-631-6	Sequence 6, Appl1
49	167.5	9.1	228	1	US-08-442-248-4	Sequence 4, Appl1
50	167.5	9.1	228	1	US-08-440-815-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-214-631-3
Sequence 3, Application US/09214631
Patent No. 6413730
GENERAL INFORMATION:
APPLICANT: Hollandu, Sacha
APPLICANT: Mbamalu, Gerardine
APPLICANT: Pawson, Tony
TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED
TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR
FILE REFERENCE: 11757.23USWO
CURRENT APPLICATION NUMBER: US/09/214,631
EARLIER FILING DATE: 1999-03-12
EARLIER APPLICATION NUMBER: PCT/CA97/00473
EARLIER FILING DATE: 1997-07-04
EARLIER APPLICATION NUMBER: 60/021,272
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 340
TYPE: PRT
ORGANISM: Homo sapiens
US-09-214-631-3

Query Match 99.8%, Score 1846, DB 4, Length 340;
Best Local Similarity 99.7%, Pred. No. 1.1e-146;
Matches 339; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPPHSGGCGVGVGALLLLGVLGVLSGLSPYVYNSANKRFQARBGVYVYPOGDRDL 60
DB 1 MGPPHSGGCGVGVGALLLLGVLGVLSGLSPYVYNSANKRFQARBGVYVYPOGDRDL 60
QY 61 LCPRARPPGPHSSPYRYEYKYLIVGAGRCCEAPAPNLLTCDRPLDRLFTTKFOEY 120
DB 61 LCPRARPPGPHSSPYRYEYKYLIVGAGRCCEAPAPNLLTCDRPLDRLFTTKFOEY 120
QY 121 SGNLWGHFRSHHDYIYIATSDGTREGLESYQGGVCLTRKVKVILRVQSGFRGCAVPKPK 180
DB 121 SGNLWGHFRSHHDYIYIATSDGTREGLESYQGGVCLTRKVKVILRVQSGFRGCAVPKPK 180
QY 161 VSEMEREDRGAAHSLFEGKENTLPDPTSNATSGABEGLPPSPMAVAGAGATALLL 240
DB 161 VSEMEREDRGAAHSLFEGKENTLPDPTSNATSGABEGLPPSPMAVAGAGATALLL 240

QY 241 GYAGAGACWRRRRRAKPSRHPGSGRGCSIGCGGGMGPRAEPGELIALRG 300
DB 241 GYAGAGACWRRRRRAKPSRHPGSGRGCSIGCGGGMGPRAEPGELIALRG 300
QY 301 GAADPPFCPHYKESGDYGHVYIVODGPPGSPNITYKV 340
DB 301 GAADPPFCPHYKESGDYGHVYIVODGPPGSPNITYKV 340

RESULT 2

US-08-436-044-2
Sequence 2, Application US/08436044
Patent No. 5624899
GENERAL INFORMATION:
APPLICANT: Bennett, Brian D.
APPLICANT: Matthews, William
TITLE OF INVENTION: HTK LIGAND
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,044
FILING DATE: 05-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/277722
FILING DATE: 20-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 902D3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 336 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-436-044-2

Query Match 34.5%; Score 637.5; DB 1; Length 336;
Best Local Similarity 42.1%; Pred. No. 1e-45;
Matches 139; Conservative 49; Mismatches 129; Indels 13; Gaps 5;
QY 14 GALLILGVLGSLSEFYVYNSANKRFOAGGYLYPOIGDRLDLCPRARPPGPHSS 73
DB 17 GLMLVLCRTALISRYLEPIYVYNSNSKFLPGQGLVLYPOIGDKLDICPKV--DSKTV 73
QY 74 PNYEFYKLVYVGAAGRCRCEAPPAPNLLITCDPDLRFTIKFOEYSPNLMGHEFRSH 133
DB 74 GQYEVYKVVVWDKQADRCITIKENTPILNCARPDQDVFTIKFOEYSPNLMGHEFRSH 133
QY 134 DYYTIATSDGTREGLSLGGVCLTRGMVTLRVGQ--SPRGAVPRKPVSEMPMER-DR 190
DB 134 DYYTIATSDGTREGLSLGGVCLTRGMVTLRVGQ--SPRGAVPRKPVSEMPMER-DR 190
QY 191 GAHSLPEKXENLPDPTSNATSRGAEGLPPSPMPAVAGAAGLALLILGAVAGAGAMC 250
DB 194 GRSSTSPFKVPKPGSGSTDSNSAGHSGNNLGSVALFAGIASGCIIFVITITLVLL 253

QY 251 WRRRAKPSRHRPBGSGFGRGSLGLGCGGGMGPRAEPGELIALRGGAADPPFCPH 310
DB 254 KYRRRRKPSRHRPBGSTTLSTLSTLATPKRGNN-----NGSRSPVITPLR---TADSVFCH 306
QY 311 YEKVSGDYGHVYIVODGPPGSPNITYKV 340
DB 307 YEKVSGDYGHVYIVODGPPGSPNITYKV 336

RESULT 3

US-08-436-054-2
Sequence 2, Application US/08436054
Patent No. 5864020
GENERAL INFORMATION:
APPLICANT: Bennett, Brian D.
APPLICANT: Matthews, William
TITLE OF INVENTION: HTK LIGAND
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,054
FILING DATE: 05-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/277722
FILING DATE: 20-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 902D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 336 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-436-054-2

Query Match 34.5%; Score 637.5; DB 2; Length 336;
Best Local Similarity 42.1%; Pred. No. 1e-45;
Matches 139; Conservative 49; Mismatches 129; Indels 13; Gaps 5;
QY 14 GALLILGVLGSLSEFYVYNSANKRFOAGGYLYPOIGDRLDLCPRARPPGPHSS 73
DB 17 GLMLVLCRTALISRYLEPIYVYNSNSKFLPGQGLVLYPOIGDKLDICPKV--DSKTV 73
QY 74 PNYEFYKLVYVGAAGRCRCEAPPAPNLLITCDPDLRFTIKFOEYSPNLMGHEFRSH 133
DB 74 GQYEVYKVVVWDKQADRCITIKENTPILNCARPDQDVFTIKFOEYSPNLMGHEFRSH 133
QY 134 DYYTIATSDGTREGLSLGGVCLTRGMVTLRVGQ--SPRGAVPRKPVSEMPMER-DR 190
DB 134 DYYTIATSDGTREGLSLGGVCLTRGMVTLRVGQ--SPRGAVPRKPVSEMPMER-DR 190
QY 191 GAHSLPEKXENLPDPTSNATSRGAEGLPPSPMPAVAGAAGLALLILGAVAGAGAMC 250
DB 194 GRSSTSPFKVPKPGSGSTDSNSAGHSGNNLGSVALFAGIASGCIIFVITITLVLL 253
QY 251 WRRRAKPSRHRPBGSGFGRGSLGLGCGGGMGPRAEPGELIALRGGAADPPFCPH 310

Db 254 KYRRRRHAKHSPQHTTTLSTLTPKRGANN-----NSERSDVITPLR---TDSVFCCH 306
 QY 311 YEKVSGDYGHPVYIVQDGPSPNNIYKV 340
 Db 307 YEKVSGDYGHPVYIVQDGPSPNNIYKV 336

RESULT 4
 PCT-US95-08812-2
 ; Sequence 2, Application PC/TUS9508812
 ; GENERAL INFORMATION:
 ; APPLICANT: Genentech, Inc.
 ; TITLE OF INVENTION: HTK LIGAND
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: patin (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/08812
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lee, Wendy M.
 ; REGISTRATION NUMBER: 00,000
 ; REFERENCE/DOCKET NUMBER: 902PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-1994
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 336 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; PCT-US95-08812-2

Query Match 34.5%; Score 637.5; DB 5; Length 336;
 Best Local Similarity 42.1%; Pred. No. 1e-45;
 Matches 139; Conservative 49; Mismatches 129; Indels 13; Gaps 5;
 QY 14 GALLILGLVGLVSLSPYVYNSANKRFQAGGYLYVPOIGDLDLCPRRAPPGPHSS 73
 Db 17 GLVAVLCRTASRIVLEPIYVNSNSKFLBGQGLVIVPQIGDLDLICRKY---DSKIV 73
 QY 74 PNYEYKLYLVGAQGRRCAPPAVNLITCDRPDLRTFTIKQVSEYVLMKHEPESHA 133
 Db 74 GQYEVYKVVYDKQADRCITIKENTPLINCARPDQDVKFTIKQVSEYVLMKHEPESHA 133
 QY 134 DYVLIANSDFRREGLESLOGVCLTRGKMYLLRVGO--SPRGGVVPKPPSEMMER-DR 150
 Db 134 DYVLIANSDFRREGLESLOGVCLTRGKMYLLRVGO--SPRGGVVPKPPSEMMER-DR 150
 QY 191 GAASLEBPKENTLPDDPTSNATSRGAGPLPPSPMPAVAGAAGLALLLLGVAAGGAMC 250
 Db 194 GRSSFTSPFVXPNGSSHDGASAGHSGNNLLGSEVALFAGIASGCIIFIVIIITLVLL 253
 QY 251 WRRRAKSESERHPGFGFGSGSLGCGGGGMPRAABPELGIATRCGGAADPPRCPH 310
 Db 254 KYRRHRHAKHSPQHTTTLSTLTPKRGANN-----NSERSDVITPLR---TDSVFCCH 306

QY 311 YEKVSGDYGHPVYIVQDGPSPNNIYKV 340
 Db 307 YEKVSGDYGHPVYIVQDGPSPNNIYKV 336

RESULT 5
 US-08-213-403-2
 ; Sequence 2, Application US/08213403
 ; Patent No. 5512457
 ; GENERAL INFORMATION:
 ; APPLICANT: Lyman, Stewart D.
 ; APPLICANT: Beckmann, M. Patricia
 ; APPLICANT: Baum, Peter R.
 ; APPLICANT: Carpenter, Melissa
 ; TITLE OF INVENTION: No. 5512457el Cytokine Designated elk ligand
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Immunex Corporation
 ; STREET: 51 University Street
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98101
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Apple Macintosh
 ; OPERATING SYSTEM: Apple System 7.1
 ; SOFTWARE: Microsoft Word for Apple, Version 5.1a
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/213,403
 ; FILING DATE: 15-MAR-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/977,693
 ; FILING DATE: 13-NOV-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Seese, Kathryn A.
 ; REGISTRATION NUMBER: 32,172
 ; REFERENCE/DOCKET NUMBER: 2807-A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 587-0430
 ; TELEFAX: (206) 233-0644
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 346 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-213-403-2

Query Match 34.2%; Score 632; DB 1; Length 346;
 Best Local Similarity 39.5%; Pred. No. 3e-45;
 Matches 145; Conservative 48; Mismatches 116; Indels 58; Gaps 9;
 QY 8 PGAVVGAALLLLGLVGLVSL-----SLSPYVYNSANKRFQAGGYLYVPOIGDLDL 61
 Db 4 PGQVWLGKVLVAVVYVMAVCLATPLAKVLEPVSWSSINPKFLSGGLVITVPIGDKLDII 63
 QY 62 CPRAAPPGPHSEPNYEFYKLYLVGAQGRRCAPPAVNLITCDRPDLRTFTIKQVSEY 121
 Db 64 CPRAAGRP-----YEVYKLYLVREPOAAACSTVLDPVNLVYCNPEDEIRFTIKQVSEY 118
 QY 122 PNLVGHFPRSHHDYIIITSDTREGLESLOGVCLTRGKMYLLRVGSPGGAVPKRPV 181
 Db 119 PNYMGLFPGKHDIYITITSNSLEGLNREGGVORTTMYKIMKVGQDPNAVTPGQLTT 178
 QY 182 SMPMERDRGAASLE-DEKENLPDDPTSNATSRGAGPLPPSPMPAVAGAAGLA----- 236
 Db 179 SRPSKADNTVYKATQAGSGRSLSDSDGKHETVQEEKSGP-----GASGSSGDPD 231
 QY 237 -----LLLVGAAGGA-----MCWRRRAKSESERHPGFGFGSGSLG 277
 Db 232 GPFNSKVALLFAAVGAGCVIFLLIIITLVLLIKRKRRKRTIQ-----RAAALSL 282

QY 278 -----GGGGGMPREAPGELIALRGGAADPPFCPHYEKVSGDYGHPIYVODGPPQSP 333
 DB 283 STLASPKGSGGTGTPSDIIIPLR---TTENNYPCHYEKSGDYGHPIYVODGPPQSP 339
 QY 334 PNITYKV 340
 DB 340 ANITYKV 346

RESULT 6
 US-08-458-077-2

/ Sequence 2, Application US/08458077
 / Patent No. 5627267
 / GENERAL INFORMATION:
 / APPLICANT: Lyman, Stewart D.
 / APPLICANT: Beckmann, M. Patricia
 / APPLICANT: Baum, Peter R.
 / APPLICANT: Carpenter, Melissa
 / TITLE OF INVENTION: No. 5627267el Cytokine Designated elx ligand
 / NUMBER OF SEQUENCES: 2
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Immunex Corporation
 / STREET: 51 University Street
 / CITY: Seattle
 / STATE: Washington
 / COUNTRY: USA
 / ZIP: 98101
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / OPERATING SYSTEM: Apple Macintosh
 / SOFTWARE: Microsoft Word for Apple, Version 5.1a
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/458,077
 / FILING DATE: 01-JUN-1995
 / CLASSIFICATION: 5.4
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 08/213,403
 / FILING DATE: 15-MAR-1994
 / APPLICATION NUMBER: US 07/977,693
 / FILING DATE: 13-NOV-1992
 / CLASSIFICATION: 5.14
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Seese, Kathryn A.
 / REGISTRATION NUMBER: 32,172
 / REFERENCE/DOCKET NUMBER: 2807-A
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (206) 587-0430
 / TELEFAX: (206) 233-0644
 / INFORMATION FOR SEQ ID NO: 2:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 346 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear
 / MOLECULE TYPE: protein
 / US-08-458-077-2

Query Match 34.2%; Score 632; DB 1; Length 346;
 Best Local Similarity 39.5%; Pred. No. 3e-45; Indels 58; Gaps 9;
 Matches 145; Conservative 48; Mismatches 116;

QY 8 PGGRVAGALLLLGLVGLVSGL-----SLPEVYVNSANKRFOAEGGYLVYPOIGDRDL 61
 DB 4 PGGRWIGKMLVAVVWVVALCRLATPLAKNLEPVWSLSINPKFSGKGLVIYPRIGKLDII 63
 QY 62 CPAPAPPGPHSSPNYEFYLYLVGAQGRCEAPAPNLLICDDPDDDLRTTFIFQES 121
 DB 64 CPAAEAGRP-----YERYKLYVRPQAAACSTVLDPNVLVTCNRPESGIRFTIKFQES 118
 QY 122 PNIMGHFRSHDYIIATSDGTREGLESIQGVCLTRGKYLRLRGVSGPRGAVPRPV 181
 DB 119 PNMGLEFKKHHDYIITSTNSLSLEGRREGVCRTRMKIMKYGQDPNNAVTPTEQLTT 178

QY 182 SEMPERDRGAASLE-PGKENLPDPTSNATSRGAEPLPPSPMPAVAGAGLA----- 236
 DB 179 SRPSKADNTVKNALQAPSGRSLGSDSKHETVNGEKSQP-----GAGSGSSGDD 231
 QY 237 -----LLILGVAGAGA-----NCWRRRRAKPSBSRIRPGSGSLGL 277
 DB 232 GFNSKVAFPAVAGACVIFLLIIFLVLLKLRGRHRKHQQ-----RAAALSL 282
 QY 278 -----GGGGGMPREAPGELIALRGGAADPPFCPHYEKVSGDYGHPIYVODGPPQSP 333
 DB 283 STLASPKGSGGTGTPSDIIIPLR---TTENNYPCHYEKSGDYGHPIYVODGPPQSP 339
 QY 334 PNITYKV 340
 DB 340 ANITYKV 346

RESULT 7

US-08-460-741-2
 / Sequence 2, Application US/08460741
 / Patent No. 5670625
 / GENERAL INFORMATION:
 / APPLICANT: Lyman, Stewart D.
 / APPLICANT: Beckmann, M. Patricia
 / APPLICANT: Baum, Peter R.
 / APPLICANT: Carpenter, Melissa
 / TITLE OF INVENTION: No. 5670625el Cytokine Designated elx ligand
 / NUMBER OF SEQUENCES: 2
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Immunex Corporation
 / STREET: 51 University Street
 / CITY: Seattle
 / STATE: Washington
 / COUNTRY: USA
 / ZIP: 98101
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / OPERATING SYSTEM: Apple Macintosh
 / SOFTWARE: Microsoft Word for Apple, Version 5.1a
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/460,741
 / FILING DATE: 02-JUN-1995
 / CLASSIFICATION: 530
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 08/213,403
 / FILING DATE: 15-MAR-1994
 / APPLICATION NUMBER: US 07/977,693
 / FILING DATE: 13-NOV-1992
 / CLASSIFICATION: 530
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Seese, Kathryn A.
 / REGISTRATION NUMBER: 32,172
 / REFERENCE/DOCKET NUMBER: 2807-A
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (206) 587-0430
 / TELEFAX: (206) 233-0644
 / INFORMATION FOR SEQ ID NO: 2:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 346 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear
 / MOLECULE TYPE: protein
 / US-08-460-741-2

Query Match 34.2%; Score 632; DB 1; Length 346;
 Best Local Similarity 39.5%; Pred. No. 3e-45; Indels 58; Gaps 9;
 Matches 145; Conservative 48; Mismatches 116;

QY 8 PGGRVAGALLLLGLVGLVSGL-----SLPEVYVNSANKRFOAEGGYLVYPOIGDRDL 61
 DB 4 PGGRWIGKMLVAVVWVVALCRLATPLAKNLEPVWSLSINPKFSGKGLVIYPRIGKLDII 63

QY 62 CPRAAPPGHSSPNVEFYKLYLVGAQGRCEAPAPNLLTCDPDLRFTTKFOEYS 121
 Db 64 CPRAEAGRP-----YERYKLYLVPEQAAACSTVLDPNVLTGCRPEQIRFTTKFOEFS 118
 QY 122 PNLGHEFRSHDYYIATSDGTREGLESIOGGVCLTRGKTVLLRVGQSPRGAVPRKPV 181
 Db 119 PNVGLEGFKHHDYITLSVNSLSBGLNREGVCRTKIMKYGQDPNAVTEQQLTT 178
 QY 182 SEMEMERDRGAHSLP-PGKENTPGDPTSNATSRGAEGLPPPSMPAVAGAAGLA----- 236
 Db 179 SRPSKADNTVKAATQAPGSRGSLGSDGKHETVNOEKSGP-----GASGSSGDPD 231
 QY 237 -----LLLVGAAGAG-----MCKRRRAKPSRHRPGPSFGRGSLGL 277
 Db 232 GFENSKVALPAVAGAGCVTLITITVLLTKLRKRRHRTQO-----RAALSL 282
 QY 278 ----GGGGMGPREAPFEGELIALRGGAADPPFCPHYEKVSDYGHPIYIVODGPPQSP 333
 Db 283 STLASPKGGSGTAGTEPSDIIIPLR---TTENNYCPHYEKVSDYGHPIYIVQEMPQSP 339
 QY 334 PNITYKV 340
 Db 340 ANITYKV 346

RESULT 8

US-08-747-240-2
 / Sequence 2, Application US/08747240
 / Patent No. 5728813
 / GENERAL INFORMATION:
 / APPLICANT: Lyman, Stewart D.
 / APPLICANT: Beckmann, M. Patricia
 / APPLICANT: Baum, Peter R.
 / APPLICANT: Carpenter, Melissa
 / TITLE OF INVENTION: No. 5728813el Cytokine Designated alk ligand
 / NUMBER OF SEQUENCES: 2
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Immunex Corporation
 / STREET: 51 University Street
 / CITY: Seattle
 / STATE: Washington
 / COUNTRY: USA
 / ZIP: 98101
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: Apple Macintosh
 / OPERATING SYSTEM: Apple System 7.1
 / SOFTWARE: Microsoft Word for Apple, Version 5.1a
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/747,240
 / FILING DATE: 12-NOV-1996
 / CLASSIFICATION: 435
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 08/213,403
 / FILING DATE: 15-MAR-1994
 / APPLICATION NUMBER: US 07/977,693
 / FILING DATE: 13-NOV-1992
 / CLASSIFICATION: 435
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Seese, Kathryn A.
 / REGISTRATION NUMBER: 32,172
 / REFERENCE/DOCKET NUMBER: 2807-A
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (206) 587-0430
 / TELEFAX: (206) 233-0644
 / INFORMATION FOR SEQ ID NO: 2:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 346 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear
 / MOLECULE TYPE: protein
 / US-08-747-240-2

Query Match 34.2%; Score 632; DB 1; Length 346;
 Best Local Similarity 39.5%; Pred. No. 3e-45;
 Matches 145; Conservative 48; Mismatches 116; Indels 58; Gaps 9;

QY 8 PGVRVGLALLGLVGVSGP-----SLPEVYNSANKRFOAGGYLYPQIDRDL 61
 Db 4 PGQWLGKVLVAMVYMLCRLATPLACNLSPVSKSSINPFLSGKGLVYIPKIDKDI 63
 QY 62 CPRAAPPGHSSPNVEFYKLYLVGAQGRCEAPAPNLLTCDPDLRFTTKFOEYS 121
 Db 64 CPRAEAGRP-----YERYKLYLVPEQAAACSTVLDPNVLTGCRPEQIRFTTKFOEFS 118
 QY 122 PNLGHEFRSHDYYIATSDGTREGLESIOGGVCLTRGKTVLLRVGQSPRGAVPRKPV 181
 Db 119 PNVGLEGFKHHDYITLSVNSLSBGLNREGVCRTKIMKYGQDPNAVTEQQLTT 178
 QY 182 SEMEMERDRGAHSLP-PGKENTPGDPTSNATSRGAEGLPPPSMPAVAGAAGLA----- 236
 Db 179 SRPSKADNTVKAATQAPGSRGSLGSDGKHETVNOEKSGP-----GASGSSGDPD 231
 QY 237 -----LLLVGAAGAG-----MCKRRRAKPSRHRPGPSFGRGSLGL 277
 Db 232 GFENSKVALPAVAGAGCVTLITITVLLTKLRKRRHRTQO-----RAALSL 282
 QY 278 ----GGGGMGPREAPFEGELIALRGGAADPPFCPHYEKVSDYGHPIYIVODGPPQSP 333
 Db 283 STLASPKGGSGTAGTEPSDIIIPLR---TTENNYCPHYEKVSDYGHPIYIVQEMPQSP 339
 QY 334 PNITYKV 340
 Db 340 ANITYKV 346

RESULT 9

US-08-239-567-6
 / Sequence 6, Application US/08299567
 / Patent No. 5747033
 / GENERAL INFORMATION:
 / APPLICANT: Davis, et al.
 / TITLE OF INVENTION: METHOD OF ENHANCING THE BIOLOGICAL
 / TITLE OF INVENTION: ACTIVITY OF EPH FAMILY LIGANDS
 / NUMBER OF SEQUENCES: 8
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Regeneron Pharmaceuticals, Inc.
 / STREET: 777 Old Saw Mill River Road
 / CITY: Tarrytown
 / STATE: New York
 / COUNTRY: U.S.A.
 / ZIP: 10591-6707
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: PatentIn Release #1.0, Version #1.30
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/299,567
 / FILING DATE: 01-SEP-1994
 / CLASSIFICATION: 435
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Kempster, Gail M.
 / REGISTRATION NUMBER: 32,143
 / REFERENCE/DOCKET NUMBER: REG 290
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 914-345-7400
 / TELEFAX: 914-345-7721
 / INFORMATION FOR SEQ ID NO: 6:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 346 amino acids
 / TYPE: amino acid
 / STRANDEDNESS:
 / TOPOLOGY: unknown
 / MOLECULE TYPE: protein
 / US-08-239-567-6

US-08-299-567-6

Query Match	34.2%;	Score 632;	DB 1;	Length 346;
Best Local Similarity	39.5%;	Pred. No. 3e-45;		
Matches 145; Conservative	48;	Mismatches 116;	Indels 58;	Gaps 9

```

QY      8 PGGRVAGMLLLGLVGLVSGL-----SLEPVYVNSANKRFQAGAGYVLPQIDRDL 61
Db      4 PGGRVGLKVLVAVVYVMAICRLATPLAKVLEFVSVSSSLNPKLSKGLVITYPKIDKDL 63

QY      62 CPRAAPPGPHSSPNVEFYKLYVGAQGRCEAPAPNLLITCDRPLDLRFTIKFOEYS 121
Db      64 CPRAAGAPR-----YEYKLYLVRPEQAASCTYLDLPRVLTQNRPEGEIRFTIKFOEFS 118

QY      122 PVLMEHRRSHHDYIIITISGTREGLSLOGGYCLTGMKVVLLRVGQSPRPGAVPRKPV 181
Db      119 PNYMLEKPKHHDIYITSTNGSDEGLENRBGVCRTFTMKIKVQGDPRNAVTPEDLTT 178

QY      182 SEMPERDARGAHSLE-IGKENLPEDDPTSNATSGAEGPLRPMPAPVAGAAGLA----- 236
Db      179 SRPSKADNTYKMAIQAPGRSGSLGDSGKAEITVQEKSGP-----GASGSSGSDPD 231

QY      237 -----LLLLVAGAAGA-----MCWRRRRAKSESRRHPGSGFGRGSLGL 277
Db      232 GFENSKVALLPAVAGAGCVTFLIIIFLVLLLKLRKRKRKTKTQ-----RAAALSL 282

QY      278 ----CGGGGCMGRREAEPELGLIARGGSAADPRPCPHYEKVSGDYGHVPYITVQDPPQSP 333
Db      283 STLASPKGSGAGTAGEPEDIIPLR-----TTENNVCPEHKEKSGDYGHVPYIVQEMPFOSP 339

QY      334 PNITYKV 340
Db      340 ANITYKV 346

```

RESULT 10
 US-08-436-044-4
 Sequence 4, Application US/08436044
 Patent No. 5624899
 GENERAL INFORMATION:
 APPLICANT: Bennett, Brian D.
 APPLICANT: Matthews, William
 TITLE OF INVENTION: HTK LIGAND
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESS: Genentech, Inc.
 STREET: 460 Point San Bruno Blvd
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: patin (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/436,044
 FILING DATE: 05-MAY-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/277722
 FILING DATE: 20-JUL-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Lee, Wendy M.
 REGISTRATION NUMBER: 00,000
 REFERENCE/DOCKET NUMBER: 902D3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/225-1994
 TELEFAX: 415/952-9881
 TELEX: 910/371-7168
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:

```

; LENGTH: 333 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-436-044-4

```

[illegible]

RESULT 11
 / Sequence 4, Application US/08436054
 / Patent No. 5864020
 / GENERAL INFORMATION:
 / APPLICANT: Bennett, Brian D.
 / APPLICANT: Matthews, William
 / TITLE OF INVENTION: 'HTK LIGAND
 / NUMBER OF SEQUENCES: 7
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Genentech, Inc.
 / STREET: 460 Point San Bruno Blvd
 / CITY: South San Francisco
 / STATE: California
 / COUNTRY: USA
 / ZIP: 94080
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 / COMMENTS: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: patin (Genentech)
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/436,054
 / FILING DATE: 05-MAY-1995
 / CLASSIFICATION: 435
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: 08/277722
 / FILING DATE: 20-JUL-1994
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Lee, Wendy M.
 / REGISTRATION NUMBER: 00,000
 / REFERENCE/DOCKET NUMBER: 902D1
 / TELEPHONE: 415/225-1994
 / TELEFAX: 415/952-9881
 / TELE: 910/371-7168
 / INFORMATION FOR SEQ ID NO: 4:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 333 amino acids

```

; TYPE: amino acid
; TOPOLOGY: linear
US-08-436-054-4
Query Match
Best Local Similarity 40.9%; Score 629.5; DB 2; Length 333;
Matches 135; Conservative 52; Mismatches 130; Indels 13; Gaps 5;

QY 14 GALLLLGLVLSGLSEPIYWNSSANKRFQAEGLVLYPQIGRLDLCPARPFGPHSS 73
Db 14 GVLMLCRATISKSYLIEPIYWNSSNKFPLPGGLVLYPQIGKDLICPKV---DSKTV 70
QY 74 PNYEFKLYLVGAQGRCEAPAPNLLTCDBPDLDLFTIKFOEYSFNLNGHEFRSH 133
Db 71 GQYEVYKVMVDKQADRCITKENTPLNCAKPDODIKFTIKFOESFNLNGLEFQKX 130
QY 134 DYIIATSDGREGLSLQGVCLTRGMKYLRLVGQ--SPRGAVPRKPVSEMPMER-DR 190
Db 131 DYIIISTNSGLELDNQGEGVCOTRAMKILMKVGDASSAGSTRNKDPTRRPELEAGTN 190
QY 191 GAASLEPGENLPDPTSNATSRGAEGPLPPSPMPAVAGAAAGLALLLVAGAAGAMC 250
Db 191 GRSSITSPFKVKNPDSSTDNSSAGSGNNILGSEVALFAGIASGCTIFIVITLVVLL 250
QY 251 WRRRAKPSERSRHPGSGFSGSLGCGGGMGPREEAPBELGIALRGGAADPFCPH 310
Db 251 KYRRRRKHSPOHTTSLSTLATPRSGNN---NGSESDIILPLR---TADSVFCPH 303
QY 311 YEKVSGDYGHPVYIVODGPPQSPENIYKYV 340
Db 304 YEKVSGDYGHPVYIVQEMPPQSPANIYKYV 333

RESULT 12
US-08-271-948-2
; Sequence 2, Application US/08271948
; Patent No. 6303769
; GENERAL INFORMATION:
; APPLICANT: Cellecti, Douglas P.
; APPLICANT: Reddy, Pranthitha
; TITLE OF INVENTION: No. 6303769e1 Cytokine Designated Lerk-5
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Apple Macintosh
; SOFTWARE: Microsoft Word, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/271,948
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2823
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-271-948-2

```

```

Query Match
Best Local Similarity 40.9%; Score 629.5; DB 4; Length 333;
Matches 135; Conservative 52; Mismatches 130; Indels 13; Gaps 5;

QY 14 GALLLLGLVLSGLSEPIYWNSSANKRFQAEGLVLYPQIGRLDLCPARPFGPHSS 73
Db 14 GVLMLCRATISKSYLIEPIYWNSSNKFPLPGGLVLYPQIGKDLICPKV---DSKTV 70
QY 74 PNYEFKLYLVGAQGRCEAPAPNLLTCDBPDLDLFTIKFOEYSFNLNGHEFRSH 133
Db 71 GQYEVYKVMVDKQADRCITKENTPLNCAKPDODIKFTIKFOESFNLNGLEFQKX 130
QY 134 DYIIATSDGREGLSLQGVCLTRGMKYLRLVGQ--SPRGAVPRKPVSEMPMER-DR 190
Db 131 DYIIISTNSGLELDNQGEGVCOTRAMKILMKVGDASSAGSTRNKDPTRRPELEAGTN 190
QY 191 GAASLEPGENLPDPTSNATSRGAEGPLPPSPMPAVAGAAAGLALLLVAGAAGAMC 250
Db 191 GRSSITSPFKVKNPDSSTDNSSAGSGNNILGSEVALFAGIASGCTIFIVITLVVLL 250
QY 251 WRRRAKPSERSRHPGSGFSGSLGCGGGMGPREEAPBELGIALRGGAADPFCPH 310
Db 251 KYRRRRKHSPOHTTSLSTLATPRSGNN---NGSESDIILPLR---TADSVFCPH 303
QY 311 YEKVSGDYGHPVYIVODGPPQSPENIYKYV 340
Db 304 YEKVSGDYGHPVYIVQEMPPQSPANIYKYV 333

RESULT 13
PCT-US95-08534-2
; Sequence 2, Application PC/TUS9508534
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; TITLE OF INVENTION: Novel Cytokine Designated Lerk-5
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Apple Macintosh
; SOFTWARE: Microsoft Word, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08534
; FILING DATE: 06-JUL-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/271,948
; FILING DATE: 08-JUL-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Anderson, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2823-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-08534-2

```

Query Match	34.0%;	Score 629.5;	DB 5;	Length 333;
Best Local Similarity	40.9%;	Pred. No. 4.6e-45;		
Matches 135;	Conservative 52;	Mismatches 130;	Indels 13;	Gaps 5

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	34.0%	Score 629.5;	DB 5;	Length 333;
Best Local Similarity	40.9%	Pred. No. 4.6e-45;		
Matches 135;	Conservative 52;	Mismatches 130;	Indels 13;	Gaps 5;

[illegible]

```

RESULT 15
US-09-214-631-4
: Sequence 4, Application US/09214631
: Patent No. 6413730
: GENERAL INFORMATION:
: APPLICANT: Holland, Sacha
: APPLICANT: Mbamalu, Geraldine
: APPLICANT: Pavson, Tony
: TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED
: TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR
: TITLE OF INVENTION: TYROSINE KINASES
: FILE REFERENCE: 1157.23USMO
: CURRENT APPLICATION NUMBER: US/09/214,631
: CURRENT FILING DATE: 1999-03-12
: EARLIER APPLICATION NUMBER: PCT/CA97/00473
: EARLIER FILING DATE: 1997-07-04
: EARLIER APPLICATION NUMBER: 60/021,272
: EARLIER FILING DATE: 1996-07-05
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 4
: LENGTH: 333
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-214-631-4

```

Query Match	33.6%	Score 622.5	DB 4	Length 333
Best Local Similarity	40.9%	Pred. 1.8e-44		
Matches 135	Conservative 54	Mismatches 128	Indels 13	Gaps 6

QY	14	GALLILGLVLGSLSTLPVTVNNSANKRFQABGGVYLYPQIGDRDLDCPRARPPGPHS	73
DB	14	GLVMTLCETALSKSLVLEPIVYNSSNSFLEPQOGLVLPQIDKDIICPKF--DSKTV	70
QY	74	PVYEFYKYLIVGAGAGRCSEAPPAENLLITCDRPLDTRFTIKFOEYSNLTWGHEHRSIH	133
DB	71	GGYEYKTYMWDKQADKCTIKKENTPLLNQAKPQDIKFTIKFGEFSNLTWGLEFQXK	130
QY	134	DYIIATSDGPRREGLESIGGVCYCLTRGKKVYLIVQO-SPRGAVYRKPVSSEMP-MER-DR	190
DB	131	DYIIISTNGSLGEGIDNEGVCQTPRAKILMKVQDQDSSASTNKKDPTRRPELAGTN	190
QY	191	GAHSLBEGKELTPDPTSNATSRQAEGLPPEPSPAPAVAGAATALLLVAGAAGANVC	250
DB	191	GRSSPTSPFKVKNPQSSITDNGSAGHSANNILGSEVALPAGIASGCIIFVILITLVALL	250
QY	251	WRRRAKSESHRPPGSGFGRCGSLGLOGGGMGREARPGSLGALNRGGAADPPCFH	310

```

Db 251 KYRRRRKHSPOHTTTLSTLSTLAPKRSKNN---NGSEPSDIIIPLR---TADSVFCH 303
QY 311 YEKSGDYGHPIYIVODGPPSPNNIYKYV 340
Db 304 YEKSGDYGHPIYIVODGPPSPNNIYKYV 333

```

RESULT 16

```

US-09-214-631-5
; Sequence 5, Application US/09214631
; Patent No. 6413730
; GENERAL INFORMATION:
; APPLICANT: Holland, Sacha
; APPLICANT: Mbamalu, Geraldine
; APPLICANT: Pawsom, Tony
; TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED
; TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR
; TITLE OF INVENTION: TYROSINE KINASES
; FILE REFERENCE: 11757.23USMO
; CURRENT APPLICATION NUMBER: US/09/214,631
; EARLIER FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: PCT/CA97/00473
; EARLIER FILING DATE: 1997-07-04
; EARLIER APPLICATION NUMBER: 60/021,272
; EARLIER FILING DATE: 1996-07-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-214-631-5

```

Query Match

```

32.9%; Score 608.5; DB 4; Length 345;
Best Local Similarity 38.7%; Pred. No. 2.7e-43;
Matches 142; Conservative 50; Mismatches 116; Indels 59; Gaps 10;

```

```

QY 8 PGYRVG---ALILLVLGLVSGL--SLPEVTWNSANKRFQAEGGVYVLPQIGRIDLI 61
Db 4 PGGMWLAKMLYAMVMAICRLATPLAKLPEVSMSSLNPKFTSGKGLVIVPKIGDLDLI 63
QY 62 CPRARPPGPHSSPYVEFYKLYLVGAQGRCEAPPAVNLTLTCDRPLDGLRFTIKFOEYS 121
Db 64 CPRLAARP-----YKYKLYVRPEQAACSTVLDPMVLTGNRPQGRFTIKFOEYS 118
QY 122 PNLWGEHFRSHDYIATSDGTREGIESLQGGVCLTRGKVKLLRVGQSPRGGAVERPKYV 181
Db 119 PNYVGLFEFKKHHDYITSTENGSLGLEENREGVCRTRTWKIKMKVQGDNPNAVTPQLTT 178
QY 182 SEMMERDRGAASLE-PGKENLPGDPTSNATSRGAEGLPPSPMAVVGAAAGLA----- 236
Db 179 SRPEKEDNTVYKATQAPGRSGSLGSDGKHETVNOEKSGP-----GASGSSGDDPD 231
QY 237 -----LLLLGVAGAGA-----MCTRRRRAPRSESRRPGPSFGRGSLGL 277
Db 232 GFENSKVALFAAVAGACVIFLLITIFVTLTKLPKHRRHTQ-----RAALSL 281
QY 278 -----GGGGGMPREAREPGLGIALRGGAADPPCFHYEKVSDYGHPIYIVODGPPSP 333
Db 282 STIASPKGSGTAGTSPSDIILPL---FTTENYCPHYEKVSDYGHPIYIVODGPPSP 338
QY 334 PNIYKYV 340
Db 339 ANIYKYV 345

```

RESULT 17

```

US-09-214-631-13
; Sequence 13, Application US/09214631
; Patent No. 6413730
; GENERAL INFORMATION:
; APPLICANT: Holland, Sacha

```

```

; APPLICANT: Mbamalu, Geraldine
; APPLICANT: Pawsom, Tony
; TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED
; TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR
; TITLE OF INVENTION: TYROSINE KINASES
; FILE REFERENCE: 11757.23USMO
; CURRENT APPLICATION NUMBER: US/09/214,631
; EARLIER FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: PCT/CA97/00473
; EARLIER FILING DATE: 1997-07-04
; EARLIER APPLICATION NUMBER: 60/021,272
; EARLIER FILING DATE: 1996-07-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-214-631-13

```

Query Match

```

26.9%; Score 498; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 8.6e-35;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 252 RRRRAKPSRHHPPGSGRGGSLGLGGGGMGPREAPGELGIALRGGAADPPFCPHY 311
Db 1 RRRRAKPSRHHPPGSGRGGSLGLGGGGMGPREAPGELGIALRGGAADPPFCPHY 60
QY 312 EKVSQDYGHPIYIVODGPPSPNNIYKYV 340
Db 61 EKVSQDYGHPIYIVODGPPSPNNIYKYV 89

```

RESULT 18

```

US-09-214-631-11
; Sequence 11, Application US/09214631
; Patent No. 6413730
; GENERAL INFORMATION:
; APPLICANT: Holland, Sacha
; APPLICANT: Mbamalu, Geraldine
; APPLICANT: Pawsom, Tony
; TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED
; TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR
; TITLE OF INVENTION: TYROSINE KINASES
; FILE REFERENCE: 11757.23USMO
; CURRENT APPLICATION NUMBER: US/09/214,631
; EARLIER FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: PCT/CA97/00473
; EARLIER FILING DATE: 1997-07-04
; EARLIER APPLICATION NUMBER: 60/021,272
; EARLIER FILING DATE: 1996-07-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-214-631-11

```

Query Match

```

10.9%; Score 201; DB 4; Length 82;
Best Local Similarity 47.3%; Pred. No. 5.1e-10;
Matches 44; Conservative 9; Mismatches 24; Indels 16; Gaps 3;

```

```

QY 252 RRRRAKPSRHHPPGSGRGGSLGL-----GGGGMGPREAREPGLGIALRGGAADPPF 307
Db 2 RRRRAKPSRHHPPGSGRGGSLGL-----RAALSLSTLASPKGSGTAGTSPSDIILPLR---TTENNY 49
QY 308 CPHYEKVSDYGHPIYIVODGPPSPNNIYKYV 340
Db 50 CPHYEKVSDYGHPIYIVODGPPSPNNIYKYV 82

```

RESULT 19

US-09-214-631-12
Sequence 12, Application US/09214631
Patent No. 6413730
GENERAL INFORMATION:
APPLICANT: Holland, Sacha
APPLICANT: Mbamalu, Geraldine
APPLICANT: Pawson, Tony
TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED
BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR
TITLE OF INVENTION: TYROSINE KINASES
FILE REFERENCE: 11757,23USWO
CURRENT FILING DATE: 1999-03-12
EARLIER APPLICATION NUMBER: PCT/CA97/00473
EARLIER FILING DATE: 1997-07-04
EARLIER APPLICATION NUMBER: 60/021,272
EARLIER FILING DATE: 1996-07-05
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO: 12
LENGTH: 82
TYPE: PRT
ORGANISM: Homo sapiens
US-09-214-631-12

Query Match 10.8%; Score 200.5; DB 4; Length 82;
Best Local Similarity 71.7%; Pred. No. 5,6e-10;
Matches 38; Conservative 4; Mismatches 8; Indels 3; Gaps 1;

QY 288 ARBEGELIGLRGGADPFCPEHYKSGDYGHAPYIVIQDGPSPPIYKYV 340
Db 33 SEPSDIIPLR---TADSVCFPEHYKSGDYGHAPYIVQEMPPSPPIYKYV 82

RESULT 20
US-08-299-567-5
Sequence 5, Application US/08299567
Patent No. 5747033
GENERAL INFORMATION:
APPLICANT: Davis, et al.
TITLE OF INVENTION: METHOD OF ENHANCING THE BIOLOGICAL
TITLE OF INVENTION: ACTIVITY OF EPH FAMILY LIGANDS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: New York
COUNTRY: U.S.A.
ZIP: 10591-6707
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/299,567
FILING DATE: 01-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kempster, Gail M.
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 290
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 234 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein

US-08-299-567-5

Query Match 9.7%; Score 179; DB 1; Length 234;
Best Local Similarity 28.4%; Pred. No. 1,3e-07;
Matches 64; Conservative 22; Mismatches 83; Indels 56; Gaps 11;

QY 7 GPGGVVALLLLIGLVGSLSEPVYNSANKRPQAEQVLYLPQIDRLDLCP--R 64
Db 24 GPG-----GALG-----NRHVVNNSNQHLLRE-GTVGVNVVDYLDIYCPHYN 67
QY 65 ARPPGHSSPNEFYKLYLVGAGRCRCEAPAPNILLTCDRBDL---DLRFTIKFOEYS 121
Db 68 SSGAGPFGGGAQVLYLVNSKNGRTCAASGFK-RMECPHAPHSIKSEKFOR 126
QY 122 PNLGHEFRSHDYIIATSDGTREGLESLOGGVCITRGMYLTLRVGSPFGGAVPRKY 181
Db 127 APLGYEFHAGHSYIYISPTNHL-----WKCLR--MKVFCASATSHSG---EKV 174
QY 182 SEMP-----MERDRGAHSL-----RGKENLP 204
Db 175 PTLPGTMDNPKINVLDEPGENPQVPLEKSIKSTSPKREHL 219

RESULT 21

US-08-240-124-2

Sequence 2, Application US/08240124
Patent No. 5516658
GENERAL INFORMATION:
APPLICANT: BECKMAN, M. P.
APPLICANT: CERRETTI, DOUGLAS P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
TITLE OF INVENTION: RECEPTOR HEX
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/240,124
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,132
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/114,426
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/109,745
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SEESE, KATHRYN A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2814-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-240-124-2

Query Match 9.7%; Score 179; DB 1; Length 238;
Best Local Similarity 28.4%; Pred. No. 1.3e-07;
Matches 65; Conservative 24; Mismatches 80; Indels 60; Gaps 12;

QY 7 GPGGVRGALLLLGLVGLVGLSLLEPYWNSANKRFOAEGGYLYVPOIGDRDLCP--R 64
DB 24 GPG-----GALG-----NRHAYWNSNQHLRRE-GYTVQVNVNDYLDICPHYN 67
QY 65 ARPGPHSP-----NYEFYKLYVGAQGRCEAPAPNLLITCDPDL---DLRFTIKF 117
DB 68 SSGVGPAGPGGAGPQVLYVWVSRNGYRCNAGQFK-RMECNRPAPHPSPITKSEKF 126
QY 118 QEYSPNLMGHEFRSHDYIATSDGTREGLESLOGGVCLTRGMKYLRLVQSPRGAVP 177
DB 127 QRYSAFSLGYEFHAGHEHYIISTPTNLH-----WKCLR-MKVYVCCASTSHSG--- 174
QY 178 RKPVSMP-----MERDGAASLE-----PKENLP 204
DB 175 EKPYPTLPQFTMGPNVKINVLDEFGENPQVPLEKSIKSTSPKREHLP 223

RESULT 22
US-08-453-943-2
Sequence 2, Application US/08453943
Patent No. 5738844
GENERAL INFORMATION:
APPLICANT: BECKMANN, M. P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,943
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/240,124
FILING DATE: 09-MAY-1994
APPLICATION NUMBER: US 08/161,132
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/114,426
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/109,745
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SEESE, KATHRYN A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2814-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-453-943-2
Query Match 9.7%; Score 179; DB 1; Length 238;
Best Local Similarity 28.4%; Pred. No. 1.3e-07;
Matches 65; Conservative 24; Mismatches 80; Indels 60; Gaps 12;

QY 7 GPGGVRGALLLLGLVGLVGLSLLEPYWNSANKRFOAEGGYLYVPOIGDRDLCP--R 64
DB 24 GPG-----GALG-----NRHAYWNSNQHLRRE-GYTVQVNVNDYLDICPHYN 67
QY 65 ARPGPHSP-----NYEFYKLYVGAQGRCEAPAPNLLITCDPDL---DLRFTIKF 117
DB 68 SSGVGPAGPGGAGPQVLYVWVSRNGYRCNAGQFK-RMECNRPAPHPSPITKSEKF 126
QY 118 QEYSPNLMGHEFRSHDYIATSDGTREGLESLOGGVCLTRGMKYLRLVQSPRGAVP 177
DB 127 QRYSAFSLGYEFHAGHEHYIISTPTNLH-----WKCLR-MKVYVCCASTSHSG--- 174
QY 178 RKPVSMP-----MERDGAASLE-----PKENLP 204
DB 175 EKPYPTLPQFTMGPNVKINVLDEFGENPQVPLEKSIKSTSPKREHLP 223

RESULT 23
US-09-057-121-2
Sequence 2, Application US/09057121
Patent No. 5969110
GENERAL INFORMATION:
APPLICANT: BECKMANN, M. P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,121
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/240,124
FILING DATE:
APPLICATION NUMBER: US 08/161,132
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/114,426
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/109,745
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SEESE, KATHRYN A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2814-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-09-057-121-2

Query Match 9.7%; Score 179; DB 2; Length 238;
Best Local Similarity 28.4%; Pred. No. 1.3e-07;
Matches 65; Conservative 24; Mismatches 80; Indels 60; Gaps 12;

QY 7 GPGGVRGALLLIGVIGLSLEPYVNSANKRFQAGGYLYPQIGDRDLCP--R 64
DB 24 GPG-----GALG-----NRHAVYNNSSNOHLRRE-GYTVQVNVNDYDIDYCPHYN 67
QY 65 ARPPGPHSSP-----NYEFKLYLVGAQGRCEAPAPNLLITCDRPLD---DLRFTIKF 117
DB 68 SSGVGEGAGPFGGAGAQYLYVNSRNGYRTCNASQGR-RNECNRPAPHPSPIKSEKF 126
QY 118 QEYSPNLWGHEFRSHDYIYIATSDTREGLSLQGGVCLTRGMKVLRLVGQSPRGAVP 177
DB 127 QRYSAFSLGYEFHAGHEYYIISTPTNHL-----WKCLR--MKVFCVCASTSHSG--- 174
QY 178 RKPVSMP-----MERDRGAHSL-----PKENLP 204
DB 175 EKVPPTLPQFTMGPNVKINVLDFEGENPQVPLKESISGTSFKREHLP 223

RESULT 24
US-09-358-734-2

Sequence 2, Application US/09358734
Patent No. 6274117
GENERAL INFORMATION:
APPLICANT: BECKMANN, M. P.
APPLICANT: CERRETTI, DOUGLAS P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
TITLE OF INVENTION: RECEPTOR HEK
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/358,734
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/240,124
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/114,426
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/109,745
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SESE, KATHRYN A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2814-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-09-358-734-2

Query Match 9.7%; Score 179; DB 4; Length 238;
Best Local Similarity 28.4%; Pred. No. 1.3e-07;
Matches 65; Conservative 24; Mismatches 80; Indels 60; Gaps 12;

QY 7 GPGGVRGALLLIGVIGLSLEPYVNSANKRFQAGGYLYPQIGDRDLCP--R 64
DB 24 GPG-----GALG-----NRHAVYNNSSNOHLRRE-GYTVQVNVNDYDIDYCPHYN 67
QY 65 ARPPGPHSSP-----NYEFKLYLVGAQGRCEAPAPNLLITCDRPLD---DLRFTIKF 117
DB 68 SSGVGEGAGPFGGAGAQYLYVNSRNGYRTCNASQGR-RNECNRPAPHPSPIKSEKF 126
QY 118 QEYSPNLWGHEFRSHDYIYIATSDTREGLSLQGGVCLTRGMKVLRLVGQSPRGAVP 177
DB 127 QRYSAFSLGYEFHAGHEYYIISTPTNHL-----WKCLR--MKVFCVCASTSHSG--- 174
QY 178 RKPVSMP-----MERDRGAHSL-----PKENLP 204
DB 175 EKVPPTLPQFTMGPNVKINVLDFEGENPQVPLKESISGTSFKREHLP 223

RESULT 25

US-08-299-567-7
Sequence 7, Application US/08299567
Patent No. 5747033
GENERAL INFORMATION:
APPLICANT: Davis, et al.
TITLE OF INVENTION: METHOD OF ENHANCING THE BIOLOGICAL
TITLE OF INVENTION: ACTIVITY OF EPH FAMILY LIGANDS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: New York
COUNTRY: U.S.A.
ZIP: 10591-6707
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/299,567
FILING DATE: 01-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kempler, Gail M.
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 290
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
INFORMATION FOR SEQ. ID NO. 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 135 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-299-567-7
Query Match 9.5%; Score 176.5; DB 1; Length 135;
Best Local Similarity 31.9%; Pred. No. 1.1e-07;
Matches 43; Conservative 14; Mismatches 73; Indels 5; Gaps 2;

QY 33 VYVNSANKRFQAGGYLYPQIGDRDLCP--RARPPEHSSPNVEFYKLYLVGAQGR 90
DB 1 VYVNSNPKFXRXEGYTYVXNDYDIDICPHYEXXXXXXXAGXXECYLYLVXEXEYX 60
QY 91 RCEAPAPNLLITCDRP---DLDRFTIKFQEYSPNLWGHEFRSHDYIYIATSDTREG 147

```
Db 61 XXXXXVWQCNRRXAHXPIKSEKFORSPFXLGKFGXHYIYITXXXPXK 120
QY 148 LESLOGVCLTRGMK 162
Db 121 LEXRXXXXXCLRXMK 135

RESULT 26
US-09-609-324A-2
; Sequence 2, Application US/09609324A
; Patent No. RE37582
; GENERAL INFORMATION:
; APPLICANT: CERRETTI, Douglas P.
; TITLE OF INVENTION: CYTOKINE DESIGNATED LERK-6
; FILE REFERENCE: A7772
; CURRENT APPLICATION NUMBER: US/09/609,324A
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 08/920,440
; PRIOR FILING DATE: 1997-08-29
; PRIOR APPLICATION NUMBER: 08/538,709
; PRIOR FILING DATE: 1995-10-03
; PRIOR APPLICATION NUMBER: 08/318,393
; PRIOR FILING DATE: 1994-10-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 184
; TYPE: PRT
; ORGANISM: LERK-6
US-09-609-324A-2

Query Match
Best Local Similarity 29.3%; Score 176; DB 1; Length 184;
Matches 58; Conservative 19; Mismatches 69; Indels 52; Gaps 7;

QY 33 VYNSANKRFQAE-----GGVLYLPQIGDLDLCPRARPPGPHSSPNVEFYKLYVGA 87
Db 10 VYNSRNPFRQVSAVGDDGGYTVESINDYLDYCPHYGAPLP-PAERMERIILVMNGE 68
QY 88 QGRCEAPPAFNNLLTCDRPLD---DLRTIKFOEYSPNLMGHEFRSHHDYIITSDGT 144
Db 69 GHASCDHRQGRGFRKWECONRPAAPGGLKFESEKQLFTPSLGFEPFRGHYYIISATP-- 126
QY 145 REGLESLOGVCLTRGMKYLIRVQSPRGAVPRKPVSEMPMERDRGAHSLBPGKENLP 204
Db 127 ----PNLVDRPCLR--LKYVYR-----PMTETLY 149
QY 205 GDP-----TSNATSRGAE 218
Db 150 EAPEPIFTSNSSCSGLGG 167

RESULT 27
US-08-920-440B-2
; Sequence 2, Application US/08920440B
; Patent No. 5919905
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; TITLE OF INVENTION: Cytokine Designated LERK-6
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: System 7.6
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,440B
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C.
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2826-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 184 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-920-440B-2

Query Match
Best Local Similarity 29.3%; Score 176; DB 2; Length 184;
Matches 58; Conservative 19; Mismatches 69; Indels 52; Gaps 7;

QY 33 VYNSANKRFQAE-----GGVLYLPQIGDLDLCPRARPPGPHSSPNVEFYKLYVGA 87
Db 10 VYNSRNPFRQVSAVGDDGGYTVESINDYLDYCPHYGAPLP-PAERMERIILVMNGE 68
QY 88 QGRCEAPPAFNNLLTCDRPLD---DLRTIKFOEYSPNLMGHEFRSHHDYIITSDGT 144
Db 69 GHASCDHRQGRGFRKWECONRPAAPGGLKFESEKQLFTPSLGFEPFRGHYYIISATP-- 126
QY 145 REGLESLOGVCLTRGMKYLIRVQSPRGAVPRKPVSEMPMERDRGAHSLBPGKENLP 204
Db 127 ----PNLVDRPCLR--LKYVYR-----PMTETLY 149
QY 205 GDP-----TSNATSRGAE 218
Db 150 EAPEPIFTSNSSCSGLGG 167

RESULT 28
US-09-173-492-2
; Sequence 2, Application US/09173492
; Patent No. 6194172
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; TITLE OF INVENTION: Cytokine Designated LERK-6
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: System 7.6
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/173,492
; CLASSIFICATION:
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,440
; FILING DATE: 29-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C.
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2826-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
```

TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 184 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-173-492-2

Query Match 9.5%; Score 176; DB 4; Length 184;
Best Local Similarity 29.3%; Pred. No. 1.7e-07;
Matches 58; Conservative 19; Mismatches 69; Indels 52; Gaps 7;

QY 33 VYVNSANKRFOAE-----GGYLYPQIGRDLDCPRAPPGPHSSPNVEFYKLYVGGA 87
DB 10 VYVNSNRPFOYSAVGGGTYVEVSINDYDLYCPHYGAPLP-PAEMERYILMVNGE 68
QY 88 QGRCEAPPAPNLLTCDPDL--DLRFTIKFOEYSPNLWGHEFRSHDYIATSOGT 144
DB 69 GHASCDHRQGRGKWEKCNRPAPAGPLKFSKFLFTFSLGFERRPGHEYYIATP-- 126
QY 145 REGLSLOGGCVLTGKMKYLIVGSPRGAVPRKPVSEMERDRGAHSLPEKKNLP 204
DB 127 ---PMLVDRPCLR--LKYVYR-----PTNETLY 149
QY 205 GDP---TSNATSRGAEG 218
DB 150 EAPPEFTSNSSCSGLGG 167

RESULT 29

US-09-173-133-2
Sequence 2, Application US/09173133

PATENT No. 6232447
GENERAL INFORMATION:
APPLICANT: Ceretti, Douglas P.
TITLE OF INVENTION: Cytokine Designated LERK-6
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: System 7.6
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/173,133
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/920,440
FILING DATE: 29-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Henry, Janis C.
REGISTRATION NUMBER: 34,347
REFERENCE/DOCKET NUMBER: 2826-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 184 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-173-133-2

Query Match 9.5%; Score 176; DB 4; Length 184;

Best Local Similarity 29.3%; Pred. No. 1.7e-07;
Matches 58; Conservative 19; Mismatches 69; Indels 52; Gaps 7;

QY 33 VYVNSANKRFOAE-----GGYLYPQIGRDLDCPRAPPGPHSSPNVEFYKLYVGGA 87
DB 10 VYVNSNRPFOYSAVGGGTYVEVSINDYDLYCPHYGAPLP-PAEMERYILMVNGE 68
QY 88 QGRCEAPPAPNLLTCDPDL--DLRFTIKFOEYSPNLWGHEFRSHDYIATSOGT 144
DB 69 GHASCDHRQGRGKWEKCNRPAPAGPLKFSKFLFTFSLGFERRPGHEYYIATP-- 126
QY 145 REGLSLOGGCVLTGKMKYLIVGSPRGAVPRKPVSEMERDRGAHSLPEKKNLP 204
DB 127 ---PMLVDRPCLR--LKYVYR-----PTNETLY 149
QY 205 GDP---TSNATSRGAEG 218
DB 150 EAPPEFTSNSSCSGLGG 167

RESULT 30

US-09-165-533-2
Sequence 2, Application US/09165533

PATENT No. 6268482
GENERAL INFORMATION:
APPLICANT: Ceretti, Douglas P.
TITLE OF INVENTION: Cytokine Designated as LERK-6
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/165,533
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/538,709
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Yalaska, Stephen L.
REGISTRATION NUMBER: 32,655
REFERENCE/DOCKET NUMBER: 2826
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 184 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-165-533-2

Query Match 9.5%; Score 176; DB 4; Length 184;
Best Local Similarity 29.3%; Pred. No. 1.7e-07;
Matches 58; Conservative 19; Mismatches 69; Indels 52; Gaps 7;

QY 33 VYVNSANKRFOAE-----GGYLYPQIGRDLDCPRAPPGPHSSPNVEFYKLYVGGA 87
DB 10 VYVNSNRPFOYSAVGGGTYVEVSINDYDLYCPHYGAPLP-PAEMERYILMVNGE 68
QY 88 QGRCEAPPAPNLLTCDPDL--DLRFTIKFOEYSPNLWGHEFRSHDYIATSOGT 144
DB 69 GHASCDHRQGRGKWEKCNRPAPAGPLKFSKFLFTFSLGFERRPGHEYYIATP-- 126

QY 145 REGSELOGVCVLTGMYLTLRVGSPRGCAVPRKPVSEWEMERDRGAHSLGKKNLP 204
 : : : :
Db 127 ----PNLVDRPCLR--LKVYVR-----PTNLTLY 149

```

RESULT 31
PCT-US95-12779-2
; Sequence 2, Application PC/TUS9512779
; GENERAL INFORMATION:
; APPLICANT: Cellect, Douglas P.
; TITLE OF INVENTION: Cytokine Designated LERK-6
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/12779
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Malaska, Stephen L.
; REGISTRATION NUMBER: 32,655
; REFERENCE/DOCKET NUMBER: 2826-NO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ. ID NO. 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 184 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-12779-2

Query Match 9.5%, Score 176, DB 5, Length 184,
Best Local Similarity 29.3%, Pred. No. 1.7e-07,
Matches 58, Conservative 19, Mismatches 69, Indels 52, Gaps 7

QY 33 VYNNANKRRQAE-----GGYVLYPOIGRLDLPAPRPPGPHSSPYVEFYKLYLVGA 87
Db 10 VYNNRNPFRQVSAVGGGGYTVFVGINYLDIYCPHGAPLP-PAEMEMYLIVYNGE 68
QY 88 QGRGCEAPPAENLLITDPRDL---DLRTITFOEYSPMLMHEFRSHDYITATSDGT 144
Db 69 GHASCDHRQGRGFKMEKNRPAPAGPLKSEKRFOLFPPSLGFEPFRPEHEYYISAP-- 126
QY 145 REGLESIQGGVCLTRGKVLRYGQSPRGAAVPRKPVSEMPYERDRGAASHLEPGKENTLP 204
Db 127 -----PNLVDRPCLR--LKYVVR-----PNTNLTLY 149
QY 205 GDP---TSNATSRGAEG 218
Db 150 EAPDFIFTSNSCGSLGG 167

RESULT 32
PCT-US95-15781-2
; Sequence 2, Application PC/TUS9515781
; GENERAL INFORMATION:

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,001
; FILING DATE: 31 MAY 1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-011CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-455-001-2
;
; Query Match
; Best Local Similarity 9.5%; Score 176; DB 1; Length 209;
; Matches 58; Conservative 19; Mismatches 69; Indels 52; Gaps 7;
;
; QY 33 VYVNSANKFQAE-----GGYVLYPQIGRDLDCPRAPPGHSSPNVEFYKLYLVGA 87
;   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
; Db 35 VYVNRNPNRFQVSAVGDGGYTVESINDYLDYCPHYGAPLP-PAERMERIILVMVNGE 93
;
; QY 88 QGRRCAPAPAPNLLTCDRPDL---DLRTTFQEYSPMLWGHFRSHDYIATSDGT 144
;   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
; Db 94 GHASCDHRGGRFRWECNRPAPGGLPKSEKFQLPFTPSLGFEPFGHEYYIATP-- 151
;
; QY 145 REGLESIQGVCLTRGMKYLIRVGQSPRGAVPRKPVSEMPERDRGAHSLPEKENLP 204
;   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
; Db 152 ---PVLVDRPCLR--LKYVVR-----FNTMETLY 174
;
; QY 205 GDP---TSNATSRGAEG 218
;   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
; Db 175 EAPEPIFTSNSSCSGLGG 192
;
; RESULT 34
; US-08-308-814-2
; Sequence 2, Application US/08308814
; Patent No. 6288476
; GENERAL INFORMATION:
; APPLICANT: Flanagan, John G.
; APPLICANT: Cheng, Hwai-Jong
; TITLE OF INVENTION: Eph Receptor Ligands, and Uses Related
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/308,814
; FILING DATE: 19-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-308-814-2
;
; Query Match
; Best Local Similarity 9.5%; Score 176; DB 4; Length 209;
; Matches 58; Conservative 19; Mismatches 69; Indels 52; Gaps 7;
;
; QY 33 VYVNSANKFQAE-----GGYVLYPQIGRDLDCPRAPPGHSSPNVEFYKLYLVGA 87
;   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
; Db 35 VYVNRNPNRFQVSAVGDGGYTVESINDYLDYCPHYGAPLP-PAERMERIILVMVNGE 93
;
; QY 88 QGRRCAPAPAPNLLTCDRPDL---DLRTTFQEYSPMLWGHFRSHDYIATSDGT 144
;   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
; Db 94 GHASCDHRGGRFRWECNRPAPGGLPKSEKFQLPFTPSLGFEPFGHEYYIATP-- 151
;
; QY 145 REGLESIQGVCLTRGMKYLIRVGQSPRGAVPRKPVSEMPERDRGAHSLPEKENLP 204
;   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
; Db 152 ---PVLVDRPCLR--LKYVVR-----FNTMETLY 174
;
; QY 205 GDP---TSNATSRGAEG 218
;   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
; Db 175 EAPEPIFTSNSSCSGLGG 192
;
; RESULT 35
; PCT-US95-11869-2
; Sequence 2, Application PC/TUS9511869
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Eph Receptor Ligands, and Uses Related
; NUMBER OF SEQUENCES: 5
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/11869
; FILING DATE: 19-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-011CP2C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; PCT-US95-11869-2
;
; Query Match
; Best Local Similarity 9.5%; Score 176; DB 5; Length 209;
; Matches 58; Conservative 19; Mismatches 69; Indels 52; Gaps 7;

```



```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-173-492-10
Query Match          9.5%; Score 175.5; DB 4; Length 213;
Best Local Similarity 36.8%; Pred. No. 2,3e-07;
Matches 43; Conservative 14; Mismatches 51; Indels 9; Gaps 3;

QY 33 VYVNSANKRFOA-----EGGYLYPQIGRLDILCPRAPPGPHSSPNVEFYKLYLVGA 87
DB 39 VYVNSNRPFHAGAGDDGGGYTVEVSINDYLDYCPHYGALP-PAERMEHYLVYVNGE 97
QY 88 QGRCEAPAPRLTLTCRDPDL---DLAFTIKFOEYSNLMGHEFRSHDYIATS 141
DB 98 GHASCDHRQGRKWEKCNRPAPGPGPLKFSKFLTFPSLGFEFRPGHEIYIAT 154

RESULT 39
US-09-173-133-10
; Sequence 10; Application US/09173133
; Patent No. 6232447
; GENERAL INFORMATION:
; APPLICANT: Cerrecti, Douglas P.
; TITLE OF INVENTION: Cytokine Designated IERK-6
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: System 7.6
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/173.133
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/920.440
; FILING DATE: 29-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C.
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2826-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 213 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-173-133-10
Query Match          9.5%; Score 175.5; DB 4; Length 213;
Best Local Similarity 36.8%; Pred. No. 2,3e-07;
Matches 43; Conservative 14; Mismatches 51; Indels 9; Gaps 3;

```

```

US-09-214-631-7
; Sequence 7; Application US/09214631
; Patent No. 6413730
; GENERAL INFORMATION:
; APPLICANT: Holland, Sacha
; APPLICANT: Mbamalu, Geraldine
; APPLICANT: Rawson, Tony
; TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED
; TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR
; FILE REFERENCE: 11757.23USWO
; CURRENT APPLICATION NUMBER: US/09/214.631
; EARLIER FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: PCT/CA97/00473
; EARLIER FILING DATE: 1997-07-04
; EARLIER APPLICATION NUMBER: 60/021.272
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-214-631-7
Query Match          9.4%; Score 174.5; DB 4; Length 233;
Best Local Similarity 27.4%; Pred. No. 3,1e-07;
Matches 60; Conservative 24; Mismatches 74; Indels 61; Gaps 11;

QY 20 GVILVSGLSLEPVYVNSANKRFOAEGGYLYPQIGRLDILCP---ARPGHSSPN 75
DB 27 GALG-----NRHAYVWNSNQHRLRE-GYTVQVNVNDYLDYCPHYNSGAGPFGGGA 80
QY 76 YEFYKLYLVG-----GAQGR---CEAPAPRLTLTCRDPDLAFTIKFOEYSNLMGH 127
DB 81 YVLYVNSRNGRTCNASQGFKEKWCNRPAPH-----SPFSEKFOYSAFSLGY 131
QY 128 EFRSHDYIATSDTREGLSLQGVCLTRGMKYLRFVGSPPGAVFRKPVSEMP-- 185
DB 132 EPHAGHEYYIYSPTHNLH-----WKCLR--MKVFVCASTSHSG--EKPVTLPGF 179
QY 186 -----MERDRGAHSLP-----PGKENLP 204
DB 180 TWGPVKNVLEDFEGENPOVPLEKXISGTSPKXELHP 218

RESULT 41
US-08-455-001-4
; Sequence 4; Application US/08455001
; Patent No. 5795734
; GENERAL INFORMATION:
; APPLICANT: Flanagan, John G.
; APPLICANT: Cheng, Hwai-Jong
; TITLE OF INVENTION: Eph Receptor Ligands, and Uses Related
; TITLE OF INVENTION: Thereeto
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,001
; FILING DATE: 31 MAY 1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:

```

NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-011CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 200 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-455-001-4

Query Match 9.3%; Score 172; DB 1; Length 200;
Best Local Similarity 35.3%; Pred. No. 4.1e-07;
Matches 49; Conservative 16; Mismatches 58; Indels 16; Gaps 5;

QY 15 ALLLGVLGVLSGLSEP-----VYMSANKRFQAGGYVLPQIGRLDLCRA 65
DB 7 ALLLAIVG-VCVMSDDPKVTSDRYAVYMSNRFH-RGDTVEVSINDYLDYCPHY 64
QY 66 RPPGHSSEPNVEFYKLYVGAQGRCEAPAPNLLTCDRPLD--DLRFTIKFOEYSP 122
DB 65 EEPFLP--AERMERVLYVMVNYEGHASCDRQKGFKEWCNRPDPSGFLKSEKQLFTF 122
QY 123 NIMGHEFRSHDYIYIATS 141
DB 123 FSLGFEFRPGHEHYIYISAS 141

RESULT 42
PCT-US95-11869-4

Sequence 4, Application PC/TUS9511869
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Eph Receptor Ligands, and Uses Related
TITLE OF INVENTION: Thereeto
NUMBER OF SEQUENCES: 5
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/11869
FILING DATE: 19-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-011CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 200 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-11869-4

Query Match 9.3%; Score 172; DB 5; Length 200;
Best Local Similarity 35.3%; Pred. No. 4.1e-07;
Matches 49; Conservative 16; Mismatches 58; Indels 16; Gaps 5;
QY 15 ALLLGVLGVLSGLSEP-----VYMSANKRFQAGGYVLPQIGRLDLCRA 65
DB 7 ALLLAIVG-VCVMSDDPKVTSDRYAVYMSNRFH-RGDTVEVSINDYLDYCPHY 64
QY 66 RPPGHSSEPNVEFYKLYVGAQGRCEAPAPNLLTCDRPLD--DLRFTIKFOEYSP 122
DB 65 EEPFLP--AERMERVLYVMVNYEGHASCDRQKGFKEWCNRPDPSGFLKSEKQLFTF 122

QY 123 NIMGHEFRSHDYIYIATS 141
DB 123 FSLGFEFRPGHEHYIYISAS 141

RESULT 43
US-08-240-124-4

Sequence 4, Application US/08240124
Patent No. 5516658
GENERAL INFORMATION:
APPLICANT: BECKMANN, M. P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
TITLE OF INVENTION: RECEPTOR HEK
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/240,124
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,132
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/114,426
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/109,745
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SEISE, KATHRYN A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2814-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-240-124-4

Query Match 9.2%; Score 169.5; DB 1; Length 201;
Best Local Similarity 29.9%; Pred. No. 6.7e-07;
Matches 66; Conservative 18; Mismatches 82; Indels 55; Gaps 10;

QY 24 LVSGSLI-EPVYMSANKRFQAGGYVLPQIGRLDLCPPARPPGHSSEPNVEFYKLY 82
DB 20 LRGSSTLRHYVYMSNPNRL-LRGDAVVELGLNDYLDIVCPHYEGGPREGP--ETFAFY 76
QY 83 LVGAQGRCEAP-PAFNLLTCDRPLDRLRFTIKFOEYSPNLMGHEFRSHDYIYIATS 141
DB 77 MWDMPGVESQAGPAPYKRWCSLFPGHVQSEKIQRTFSLGEPFLRGERTYIYISVP 136
QY 142 DGTREGLESIGGVCLTRGKYLRLVGOQPRGAVRKPVESEPMERDRGAASLEPGKE 201
DB 137 --TPE-----SSGQCL-----RLGVSVCKRKSESAPV----- 164

QY 202 NLPDPTSNATS--RGAEGPLPPSPMPAVAGAGLALLL 240
Db 165 ---GSPGSGTSGWRGCDTPSP-----LCILL 189

RESULT 44

US-08-453-943-4
Sequence 4, Application US/08453943
Patent No. 5738844
GENERAL INFORMATION:
APPLICANT: BECKMANN, M. P.
APPLICANT: CERRETTI, DOUGLAS P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
TITLE OF INVENTION: RECEPTOR HEX
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,943
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/240,124
FILING DATE: 09-MAY-1994
APPLICATION NUMBER: US 08/161,132
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/114,426
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/109,745
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SEESE, KATHRYN A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2814-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-453-943-4

Query Match 9.2%; Score 169.5; DB 1; Length 201;
Best Local Similarity 29.9%; Pred. No. 6.7e-07;
Matches 66; Conservative 18; Mismatches 82; Indels 55; Gaps 10;
QY 24 LVSGLSL-EPVYVNSANKRFQAEAGVLYLPQIGRLDLICPPARPFGHSPNTEFYKLY 82
Db 20 IIRGSSILRHVVYVNSNPRL-LRGDAVVELGNDYLDIVCPHYEGPGEPP--ETFALY 76
QY 83 LVGAAGRCRAP-PAPNLLTCDRBDLRFITKQESPVLMGHEHSHDDYIITS 141
Db 77 MVDWPGYSCQAEGRPAVYKWCSPFGHVQSEKIQRTFSLGFEFLPGETIYIISVP 136
QY 142 DGTREGLSLGGVCLTRGMKVLIRVGSPRGAAVPRKPVSEMPERDRGAHSLPEPKE 201
Db 137 --TFE-----SSGGL-----RLQVSVCKEKERSAHVP----- 164

QY 202 NLPDPTSNATS--RGAEGPLPPSPMPAVAGAGLALLL 240
Db 165 ---GSPGSGTSGWRGCDTPSP-----LCILL 189

RESULT 45

US-09-057-121-4
Sequence 4, Application US/09057121
Patent No. 5969110
GENERAL INFORMATION:
APPLICANT: BECKMANN, M. P.
APPLICANT: CERRETTI, DOUGLAS P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
TITLE OF INVENTION: RECEPTOR HEX
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,121
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/240,124
FILING DATE:
APPLICATION NUMBER: US 08/161,132
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/114,426
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/109,745
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SEESE, KATHRYN A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2814-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-057-121-4

Query Match 9.2%; Score 169.5; DB 2; Length 201;
Best Local Similarity 29.9%; Pred. No. 6.7e-07;
Matches 66; Conservative 18; Mismatches 82; Indels 55; Gaps 10;
QY 24 LVSGLSL-EPVYVNSANKRFQAEAGVLYLPQIGRLDLICPPARPFGHSPNTEFYKLY 82
Db 20 IIRGSSILRHVVYVNSNPRL-LRGDAVVELGNDYLDIVCPHYEGPGEPP--ETFALY 76
QY 83 LVGAAGRCRAP-PAPNLLTCDRBDLRFITKQESPVLMGHEHSHDDYIITS 141
Db 77 MVDWPGYSCQAEGRPAVYKWCSPFGHVQSEKIQRTFSLGFEFLPGETIYIISVP 136
QY 142 DGTREGLSLGGVCLTRGMKVLIRVGSPRGAAVPRKPVSEMPERDRGAHSLPEPKE 201
Db 137 --TFE-----SSGGL-----RLQVSVCKEKERSAHVP----- 164

Db 137 -----SSGCL-----RLQVSVCKKESAHV----- 164

QY 202 NLPDPTSNATS--RGAEGPLPPSPMPAVAGAAGLALLL 240

Db 165 ---GSPGSGTSGWRGDTSP-----LCILL 189

RESULT 46

US-09-358-734-4

/ Sequence 4, Application US/09358734

/ Patent No. 6274117

/ GENERAL INFORMATION:

/ APPLICANT: BECKMANN, M. P.

/ TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE

/ TITLE OF INVENTION: RECEPTOR HER

/ NUMBER OF SEQUENCES: 4

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: IMMUNEX CORPORATION

/ STREET: 51 UNIVERSITY STREET

/ CITY: SEATTLE

/ STATE: WASHINGTON

/ COUNTRY: USA

/ ZIP: 98101

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ OPERATING SYSTEM: Apple System 7.1

/ SOFTWARE: Microsoft Word for Apple, Version 5.1a

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/09/358,734

/ FILING DATE:

/ CLASSIFICATION:

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: 08/240,124

/ FILING DATE:

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: US 08/114,426

/ FILING DATE: 30-AUG-1993

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: US 08/109,745

/ FILING DATE: 20-AUG-1993

/ ATTORNEY/AGENT INFORMATION:

/ NAME: SEESE, KATHRYN A.

/ REGISTRATION NUMBER: 32,172

/ REFERENCE/DOCKET NUMBER: 2814-C

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (206) 587-0430

/ TELEFAX: (206) 233-0644

/ TELEX: 756822

/ INFORMATION FOR SEQ ID NO: 4:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 201 amino acids

/ TYPE: amino acid

/ TOPOLOGY: linear

/ MOLECULE TYPE: protein

US-09-358-734-4

Query Match 9.2%; Score 169.5; DB 4; Length 201;

Best Local Similarity 29.9%; Pred. No. 6.7e-07; Indels 55; Gaps 10;

Matches 66; Conservative 18; Mismatches 82;

QY 24 LVSGLSL-EPYVNSANKRFQAEAGGYLYPQIGRLDLCPRAPPGPHSGPNYFYKLY 82

Db 20 LRGSLSLHVYVYVNSNRL-LRQDAVELGLNDYLDIVCPHYGPGPEGP--ETPALY 76

QY 83 LVGAQGRRCAP-PAPYLITCORPDLRFTIKFOYSPNLMGHERSHDYIATS 141

Db 77 WMDWGVSCQAEGRAYKRWVCSLPFGHVFSERIQFTFPLGFEPLPEETYYISVP 136

QY 142 DGTREGSLQGVCLTRGMKYLRLVQSPRGAVPRKTVSEMPERDRGAHSLPQKE 201

Db 137 --TP-----SSGCL-----RLQVSVCKKESAHV----- 164

QY 202 NLPDPTSNATS--RGAEGPLPPSPMPAVAGAAGLALLL 240

Db 165 ---GSPGSGTSGWRGDTSP-----LCILL 189

RESULT 47

US-09-214-631-8

/ Sequence 8, Application US/09214631

/ Patent No. 6413730

/ GENERAL INFORMATION:

/ APPLICANT: Holland, Sacha

/ APPLICANT: Mbamalu, Geraldine

/ APPLICANT: Pawson, Tony

/ TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED

/ TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR

/ FILE REFERENCE: 11757.23USWO

/ CURRENT APPLICATION NUMBER: US/09/214,631

/ CURRENT FILING DATE: 1999-03-12

/ EARLIER APPLICATION NUMBER: PCT/CA97/00473

/ EARLIER FILING DATE: 1997-07-04

/ EARLIER APPLICATION NUMBER: 60/021,272

/ EARLIER FILING DATE: 1996-07-05

/ NUMBER OF SEQ ID NOS: 13

/ SOFTWARE: PatentIn Ver. 2.0

/ SEQ ID NO 8

/ LENGTH: 201

/ TYPE: PRT

/ ORGANISM: Homo sapiens

US-09-214-631-8

Query Match 9.1%; Score 169; DB 4; Length 201;

Best Local Similarity 29.8%; Pred. No. 7.4e-07; Indels 62; Gaps 11;

Matches 70; Conservative 19; Mismatches 84;

QY 17 LILGYL-----GLVGLSL-EPYVNSANKRFQAEAGGYLYPQIGRLDLCPRAPP 68

Db 6 LIITVMAFLGSPRLGSSLSLHVYVYVNSNRL-LRQDAVELGLNDYLDIVCPHYGPG 64

QY 69 GPHSPNVEFYLYLYVGAQRRCAP-PAPYLITCORPDLRFTIKFOYSPNLMGH 127

Db 65 GPPEGP--ETPALYVMDWGVSCQAEGRAYKRWVCSLPFGHVFSERIQFTFPLGFE 122

QY 128 EFRSHDYIATS DGTREGSLQGVCLTRGMKYLRLVQSPRGAVPRKTVSEMPER 187

Db 123 EFLPEETYYISVP--TP-----SSGCL-----RLQVSVCKKESAHV----- 155

QY 188 RDGAHSLPEKENLPDPTSNATS--RGAEGPLPPSPMPAVAGAAGLALLL 240

Db 156 RKESAHV-----GSPGSGTSGWRGDTSP-----LCILL 189

RESULT 48

US-09-214-631-6

/ Sequence 6, Application US/09214631

/ Patent No. 6413730

/ GENERAL INFORMATION:

/ APPLICANT: Holland, Sacha

/ APPLICANT: Mbamalu, Geraldine

/ APPLICANT: Pawson, Tony

/ TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED

/ TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR

/ FILE REFERENCE: 11757.23USWO

/ CURRENT APPLICATION NUMBER: US/09/214,631

/ CURRENT FILING DATE: 1999-03-12

/ EARLIER APPLICATION NUMBER: PCT/CA97/00473

/ EARLIER FILING DATE: 1997-07-04

/ EARLIER APPLICATION NUMBER: 60/021,272

/ EARLIER FILING DATE: 1996-07-05

/ NUMBER OF SEQ ID NOS: 13

/ SOFTWARE: PatentIn Ver. 2.0

Wed Feb 12 09:30:14 2003

us-10-021-121-4.ra1

Page 23

Search completed: February 11, 2003, 12:06:26
Job time : 14.6918 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 11:48:59 / Search time 24.805 Seconds
(without alignments)
1317.705 Million cell updates/sec

Title: US-10-021-121-4

Perfect score: 1850
Sequence: 1 MGPSPGSGVAVGALLLG.....PVIYQDGPQSPNNIYKV 340

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 50 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	637.5	34.5	336	2	149766	hepatoma transmembr
2	632	34.2	346	2	S46893	elk ligand - human
3	629.5	34.0	333	2	184743	hepatoma transmembr
4	613.5	33.2	345	2	148780	Stral/Ep1g2 protei
5	608.5	32.9	345	2	158406	LEK-2 - rat
6	214.5	11.6	237	2	119914	hypothetical prote
7	179	9.7	238	2	138849	LEK-3 - human
8	176	9.5	209	2	A54984	ELF-1 protein prec
9	175.5	9.5	213	2	JE0322	ephrin-A2 - human
10	170.5	9.2	228	2	A57084	repulsive axon gui
11	169.5	9.2	201	2	138850	LEK-4 - human
12	167.5	9.1	228	2	158170	LEK-7 precursor -
13	166	9.0	205	2	A36377	B61 protein precu
14	159	8.6	680	2	S31216	collagen alpha 1(I
15	154.5	8.4	1049	1	CG8078	collagen alpha 1(I
16	153.5	8.3	1670	1	CGH03B	collagen alpha 3(I
17	151.5	8.2	301	2	B31219	collagen 2 - Caeno
18	149	8.1	325	2	T32348	hypothetical prote
19	149	8.1	569	2	S42886	collagen - silkwor
20	148.5	8.0	316	2	T20497	hypothetical prote
21	148	8.0	921	2	S42617	collagen alpha 1(I
22	146.5	7.9	1315	2	A56101	collagen alpha 1(I
23	146.5	7.9	1492	2	A40333	collagen alpha 1(I
24	146.5	7.9	1774	2	B56101	collagen alpha 1(I
25	146	7.9	675	2	S20819	collagen alpha 3(I
26	145.5	7.9	305	2	T20506	hypothetical prote
27	145	7.8	674	2	S13501	collagen alpha 1(I
28	145	7.8	931	2	S13580	collagen alpha 1(I
29	144.5	7.8	438	2	S53787	collagen alpha cha

30	144	7.8	1027	2	S28774	collagen alpha cna
31	143	7.7	1747	2	A54121	collagen alpha-4 c
32	142.5	7.7	743	1	S23779	collagen alpha 1(V
33	142.5	7.7	1496	1	CGH02V	collagen alpha 2(V
34	142	7.7	744	2	S15435	collagen alpha 1(V
35	142	7.7	1029	1	S21569	collagen alpha 2(V
36	142	7.7	1763	2	S16366	collagen alpha 2(I
37	141.5	7.6	1466	1	CGH07L	collagen alpha 1(I
38	141	7.6	319	2	T32250	hypothetical prote
39	141	7.6	744	1	A34246	collagen alpha 1(V
40	141	7.6	744	1	S23298	collagen alpha 1(V
41	140.5	7.6	305	2	T30165	hypothetical prote
42	140	7.6	304	2	T22482	hypothetical prote
43	140	7.6	680	1	CGH01D	collagen alpha 1(X
44	139.5	7.5	210	2	B44984	collagen - nematod
45	139	7.5	940	2	JE0291	FB19 protein - hum
46	139	7.5	1024	2	S18251	collagen alpha 1(X
47	138.5	7.5	469	2	A24450	collagen alpha 2(V
48	138.5	7.5	1419	2	A41182	collagen alpha 1(I
49	138.5	7.5	1487	2	B41182	collagen alpha 1(I
50	138.5	7.5	1546	1	CGH02E	collagen alpha 2(X

ALIGNMENTS

RESULT 1	149766	hepatoma transmembrane kinase ligand - mouse
C/Species: Mus musculus (house mouse)		
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999		
C/Accession: 149766		
R/Bennett, B.D.; Zeigler, F.C.; Gu, Q.; Fendly, B.; Goddard, A.D.; Gillett, N.; Matthew Proc. Natl. Acad. Sci. U.S.A. 92, 1866-1870, 1995		
A/Title: Molecular cloning of a ligand for the EPH-related receptor protein-tyrosine ki		
A/Reference number: 149766; MUID:5519254; PMID:7534404		
A/Accession: 149766		
A/Status: preliminary; translated from GB/EMBL/DBJ		
A/Molecule type: mRNA		
A/Residues: 1-336 <RES>		
A/Cross-references: GB:L38847; MID:g769677; PIDN:AAC42052.1; PID:g769678		
C/Genetics:		
A/Gene: HTK		
Query Match	34.5%	Score 637.5; DB 2; Length 336;
Best Local Similarity	42.1%;	Pred. No. 1.9e-37;
Matches 139; Conservative 49; Mismatches 129; Indels 13; Gaps 5;		
QY 14 GALLIGVLTGSLSEPVYVNSANKRFQAEGVLYPQIGDRLDLCPRAPFGPHSS 73		
DB 17 GLMWLCRTAISRSIVLEPIYVNSNSKFLPGQGLVLPQIGDRLDLCPRKV--DSKIV 73		
QY 74 PNYEFYKLYVGAQGRRCAPAPNLLTCDPDLRLFTIKQEVSPNMGHFFRSH 133		
DB 74 GQYEVYKLYVMDKQADRCITKENTPILNCARPDOVKFKIKQEVSPNMGHFFRSH 133		
QY 134 DYIYIATSDGTRREGLESQGVCLTRGKVLRYGO--SPRGAVPRKPVSEMER-DR 190		
DB 134 DYIISTNSGLBEGDNQGVGVCOTRANKIMKYQDASSAGSARNNGPTRRPELAGTN 193		
QY 191 GAHSLEFGKENLPDDPSNATSRGAEGLPPSPMPAVAGAGGIALLLGVAGAGMC 250		
DB 194 GRSSITSPFVKNPSSITDGSAGHSNNLLGSEVALFAGIASGCIIFVIYIITLVLL 253		
QY 251 WRRRAKSSSRHNPQSPFGSGSLGCGGGMGRREPELGIARGGAAAPPGCH 310		
DB 254 KYRRHRHRSQHTTSLSTLATPKGKNV---NGSEBDVILPLR--TASVFCFH 306		
QY 311 YEKVSGDYGHPIYIVQDGPQSPNNIYKV 340		
DB 307 YEKVSGDYGHPIYIVQDGPQSPNNIYKV 336		

RESULT 2

S46993
 elik ligand - human
 C/Species: Homo sapiens (man)
 C/Date: 15-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 28-May-1999
 C/Accession: S46993
 R/Beckmann, M.P.; Cerretti, D.P.; Baum, P.; Vanden Bos, T.; James, L.; Farrar, T.; Kozlowski, J. 13, 3757-3762, 1994
 A/Title: Molecular characterization of a family of ligands for eph-related tyrosine kinases
 A/Reference number: S46993; MUID:94349923; PMID:8070404
 A/Accession: S46993
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-346 <BEG>
 A/Cross-references: GB:U09304; NID:G538366; PIDN:AAA53093.1; PID:G538367

Query Match 34.2%; Score 632; DB 2; Length 346;
 Best Local Similarity 39.5%; Pred. No. 4.7e-37;
 Matches 145; Conservative 48; Mismatches 116; Indels 58; Gaps 9;

```

QY 8 PGAGVGLLLGLVGLVSGT-----SLEPYVNSANKRFGAEGGYLYPOIGDRDLCLPRARPPGPHS
DB 4 PQGRMLGKVLVAVVWALCRLATPLAKNLEPVSWSLNPKFLSGKGLVITFKIGDKDIT 63
QY 62 CPRARPPGPHSSPNYEFYLYLVGAQGRRCRAPPNLLTCDRPDLLEFTIKFOEYS 121
DB 64 CPRARGR-----YEVYLYVREPEQAAGSTVLDPNVLTCKNPEQEIFTIKFOEFS 118
QY 122 PNLNGHEFSHDYITIAISDTRGLESLOGVCLTRGMKTLRVGQSPGGAIVPRKPV 181
DB 119 PVMGJEFFKHHDYITTSNGSLGLENGEGVCRTKIMKIVGODPNNAVTPEDULTT 178
QY 182 SEMPERDGAASHLE-PEKENLPDPTSNATSRGAEGPLPPSPYAVAGAAGIA---- 236
DB 179 SRPSKADVTVMATQAPSRGSLSDSGKHETVNGEKSQP-----GASGGSSGDPD 231
QY 237 -----LILLGVAGAGCA-----MCTRRRAKSESNSHPGSGRGSGISGL 277
DB 232 GFENSKVALFAAVGAGCVFLIILFVTLTLKLRKRKHQO-----RAAALSL 282
QY 278 -----GGGGGNGPREAPRGELGIALRGGAADPPCPHYEKGSDGHPYIVQDPPQSP 333
DB 283 STASKSGSGTAGTPTSDIILPLR--TTENNYCPHFKVSGDIGHPIYIVQEMPQSP 339
QY 334 PNYYKYV 340
DB 340 ANIYYKV 346

```

RESULT 3

S46993
 hepatoma transmembrane kinase ligand - human
 C/Species: Homo sapiens (man)
 C/Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 05-Nov-1999
 C/Accession: S46993
 R/Bennett, B.D.; Zeigler, F.C.; Gu, Q.; Fendly, B.; Goddard, A.D.; Gillett, N.; Matthews
 Proc. Natl. Acad. Sci. U.S.A. 92:1866-1870, 1995
 A/Title: Molecular cloning of a ligand for the EPH-related receptor protein-tyrosine kinases
 A/Reference number: 149766; MUID:95199254; PMID:7534404
 A/Accession: S46993
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-333 <RES>
 A/Cross-references: GB:I38734; NID:G769675; PIDN:AA61752.1; PID:G769676
 C/Genetic: GDB:EPG5; LERK5
 A/Genes: GDB:EPG5; LERK5
 A/Cross-references: GDB:438338; OMIM:600527
 A/Map position: 13q33-13q33

Query Match 34.0%; Score 629.5; DB 2; Length 333;
 Best Local Similarity 40.9%; Pred. No. 6.8e-37;
 Matches 135; Conservative 52; Mismatches 130; Indels 13; Gaps 5;

```

QY 14 GALLILGLVGLVGLSLEPYVNSANKRFGAEGGYLYPOIGDRDLCLPRARPPGPHS 73
DB 14 GVLNLCRTAISKSYLLEPYVNSANKRFGAEGGYLYPOIGDRDLCLPRARPPGPHS 70
QY 74 PNIEFYKLYLVGAQGRRCRAPPNLLTCDRPDLLEFTIKFOEYSNLMGHEFSRH 133
DB 71 GQVEYKVVVVDQDADRCTIKENPTPLNCAKPPDDIFITIKFOEFSNLMGHEFSRH 130
QY 134 DYTITNSDTRGLESLOGVCLTRGMKTLRVGQ--SPRGAVPRKVSSEMPER-DR 190
DB 131 DYTITNSDTRGLESLOGVCLTRGMKTLRVGQ--SPRGAVPRKVSSEMPER-DR 190
QY 191 GAASHLEPKENLPDPTSNATSRGAEGPLPPSPYAVAGAAGIALLLGVAGAGAAC 250
DB 191 GRSSTSPFKVPKVPKPGSTGNSAGHSGNNILGSEVALFAGIASGCIIFVITLTVLL 250
QY 251 WRRRRAPKPSRRPFGSGSGSGSLGAGGGGNGPREAPRGELGIALRGGAADPPCPH 310
DB 251 KYRRRRKPSRRPFGSTGNSAGHSGNNILGSEVALFAGIASGCIIFVITLTVLL 303
QY 311 YEKVSGDYGHPIYIVQDPPQSPPNYYKYV 340
DB 304 YEKVSGDYGHPIYIVQEMPQSPANYYKYV 333

```

RESULT 4

S46993
 Stral/Ep1g2 protein - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
 C/Accession: S46993; A55062; S52670
 R/Boulic, P.; Ould-Abdelgann, M.; Vicaire, S.; Garnier, J.M.; Schunauer, B.; Dolle,
 Dev. Biol. 170, 420-433, 1995
 A/Title: Efficient cloning of cDNAs of retinoic acid-responsive genes in P19 embryonal
 A/Reference number: 148780; MUID:95337533; PMID:7649373
 A/Accession: S46993
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-345 <RES>
 A/Cross-references: EMBL:248781; NID:G747858; PIDN:CAA8695.1; PID:G747859
 R/Fletcher, F.A.; Renshaw, B.; Hollingsworth, T.; Baum, P.; Lyman, S.D.; Jenkins, N.A.;
 Genomics 24, 127-132, 1994
 A/Title: Genomic organization and chromosomal localization of mouse Ep1g2, a gene encod
 A/Reference number: A55062; MUID:95203867; PMID:7896266
 A/Accession: S46993
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-345 <RES>
 A/Cross-references: GB:U07598
 R/Shao, H.; Lou, L.; Pandey, A.; Pasquale, E.B.; Dixit, V.M.
 U. Biol. Chem. 269, 26606-26609, 1994
 A/Title: cDNA cloning and characterization of a ligand for the Cdk5 receptor protein-ty
 A/Reference number: A55062; MUID:95014510; PMID:7929389
 A/Accession: S46993
 A/Status: preliminary; not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-89; 'T', '91-345 <SHA>
 A/Cross-references: GB:U12883; NID:G575928; PIDN:AAA53231.1; PID:G575929
 A/Genetic: GDB:EPG2

Query Match 33.2%; Score 613.5; DB 2; Length 345;
 Best Local Similarity 38.2%; Pred. No. 9.3e-36;
 Matches 138; Conservative 51; Mismatches 107; Indels 65; Gaps 10;

```

QY 15 ALLILGLVGLVSGT-----SLEPYVNSANKRFGAEGGYLYPOIGDRDLCLPRARPPGPHS 72
DB 15 AVVLTLCRLATPLAKNLEPVSWSLNPKFLSGKGLVITFKIGDKDITCPRARGR-- 72
QY 73 SPNIEFYKLYLVGAQGRRCRAPPNLLTCDRPDLLEFTIKFOEYSNLMGHEFSRH 132
DB 73 -----YEVYKLYLVREPEQAAGSTVLDPNVLTCKNPEQEIFTIKFOEFSNLMGHEFSRH 129

```

QY 133 HDYIITATSDGTREGESLGGVGLTRGMKVLRLVQSGSPRGAVPRKPVSEMERDRCA 192
 Db 130 HDYITSTNSGLEENREGVCRTRTKIWKVGQDP--NAVTPROLTSRPSKSDNT 188
 QY 193 AHSLEPKENLPGDPTNATSRGAEGR-----LPPSMNAVAGAGG-----LA 236
 Db 189 VKT-----ATQAPRGSGQSDGSKHETVNOEEKSGGAGGSGSDSPFNSK 236
 QY 237 LLLLGAVAGAGA-----MCMRRRRAPKPSRRHPGSGFGRGSLGL-----GG 279
 Db 237 VALFAAVAGAGCVLFLITITFLTVLLKLRKRKHQO-----RAAALSLSTLASP 287
 QY 280 GCGNGPREAPPELGIALRGGAADPPFCPHYEKVSGDYGHVPIYVQDGPSPENIYYK 339
 Db 288 KGGSGTAGTEPSDIIIFLR--TTENNYCPHYEKVSGDYGHVPIYVQDGPSPENIYYK 344
 QY 340 V 340
 Db 345 V 345

RESULT 5

158406
 LERK-2 - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 05-Nov-1999
 C/Accession: 158406
 R/Fletcher, F.A.; Carpenter, M.; Shilling, H.; Baum, P.; Ziegler, S.; Hollin
 Oncogene 9, 3241-3248, 1994
 A>Title: LERK-2, a ligand for the receptor tyrosine kinase ELK, is evolutionarily conserved
 A/Reference number: 158406; MUID:95022634; PMID:7936648
 A/Accession: 158406
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-345 <RES>
 A/Cross-references: EMBL:U07560; NID:9563118; PID:AAA53092.1; PID:9563119
 C/Genetics:
 A/Genes: Epi92

Query Match 32.9%; Score 608.5; DB 2; Length 345;
 Best Local Similarity 38.0%; Pred. No. 2.1e-35;
 Matches 137; Conservative 52; Mismatches 107; Indels 65; Gaps 10;

QY 15 ALLLLGLVGLVSG--SLSPYVWNSANKRFOAEGGYLVPOIGDRDLDCPRARPPGPHS 72
 Db 15 AMVLLTLCRLATPLAKNLEPVSWSLPKFLSGKGLVIFKIGDKDITCPRAEAGRP-- 72
 QY 73 SPNTEFYKLYLVGAQGRRCCEAPPAPNLLTCRPRDLRLFTIKFOEYSPNLMGHEFRSH 132
 Db 73 ---YEYKLYVREPEQAACSTVLDPNVLTCKNPKQGEIRFTIKFOEFSNVMGLFKKY 129
 QY 133 HDYIITATSDGTREGESLGGVGLTRGMKVLRLVQSGSPRGAVPRKPVSEMERDRCA 192
 Db 130 HDYITSTNSGLEENREGVCRTRTKIWKVGQDP--NAVTPROLTSRPSKSDNT 188
 QY 193 AHSLEPKENLPGDPTNATSRGAEGR-----LPPSMNAVAGAGG-----LA 236
 Db 189 VKT-----ATQAPRGSGQSDGSKHETVNOEEKSGGAGGSGSDSPFNSK 236
 QY 237 LLLLGAVAGAGA-----MCMRRRRAPKPSRRHPGSGFGRGSLGL-----GG 279
 Db 237 VALFAAVAGAGCVLFLITITFLTVLLKLRKRKHQO-----RAAALSLSTLASP 287
 QY 280 GCGNGPREAPPELGIALRGGAADPPFCPHYEKVSGDYGHVPIYVQDGPSPENIYYK 339
 Db 288 KGGSGTAGTEPSDIIIFLR--TTENNYCPHYEKVSGDYGHVPIYVQDGPSPENIYYK 344
 QY 340 V 340
 Db 345 V 345

RESULT 6

119914
 Hypothetical protein C43F9.8 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: 119914
 R/Mottimore, B.
 Submitted to the EMBL Data Library, November 1996
 A/Reference number: 219195
 A/Accession: 119914
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-237 <WTL>
 A/Cross-references: EMBL:Z82262; PID:CA854195.1; GSPDB:GN00022; CESP:C43F9.8
 A/Experimental source: clone C43F9
 C/Genetics:
 A/Genes: CESP:C43F9.8
 A/Map position: 4
 A/Introns: 32/2; 96/3; 214/1

Query Match 11.6%; Score 214.5; DB 2; Length 237;
 Best Local Similarity 25.7%; Pred. No. 4.7e-08;
 Matches 53; Conservative 41; Mismatches 83; Indels 29; Gaps 6;

QY 11 VRVGLLLGLVGLVSG--GLSLSPYVWNSANKRFOAEG--GYLVPOIGDRDLDCPRARPP 68
 Db 1 MGIATPILSLPPIGMAKRIPDINMISNPIFVSNDHVISVHIGDRVSRCPKSDT 60
 QY 69 GPHSPNTEFYKLYLVGAQGRRCCEAPPAPNLLTCRPRDLRLFTIKFOEYSPNLMGHE 128
 Db 61 G-----KVEYSYIYVSDSEYDHGFL--SKPRLVACQDNQININIVFRSFTPTPGFE 114
 QY 129 FRSHDYIITL-----TSDGRBSLESLGGVGLTRGMKVLRLVQ 169
 Db 115 FQPGKNYFLISSEVDALITETANQITPGTSDTBSIDKKDGLCTAKQMKIKFEVQ 174
 QY 170 SPRGAVPRKPVSEMERDRGAHS 195
 Db 175 DRGTEINPK--FAARTLKDDAHS 198

RESULT 7

138849
 LERK-3 - human
 C/Species: Homo sapiens (man)
 C/Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-Sep-1999
 C/Accession: 138849
 R/Kozlovsky, C.J.; Maraskovsky, E.; McGrew, J.T.; Vandembos, T.; Teepe, M.; Lyman, S.D.;
 Oncogene 10, 239-306, 1995
 A>Title: Ligands for the receptor tyrosine kinases hek and elk: isolation of cDNAs encod
 A/Reference number: 138849; MUID:95140419; PMID:7838529
 A/Accession: 138849
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-238 <RES>
 A/Cross-references: EMBL:U14187; NID:9642832; PID:AA50078.1; PID:9642833
 A/Genes: GDB:EPG3
 A/Cross-references: GDB:438336; OMIM:601381
 A/Map position: 1q21-q22
 A/Superfamily: axon guidance signal protein

Query Match 9.7%; Score 179; DB 2; Length 238;
 Best Local Similarity 28.4%; Pred. No. 1.4e-05;
 Matches 65; Conservative 24; Mismatches 80; Indels 60; Gaps 12;

QY 7 GPGGVAVGALLGLVGLVSG--SLSPYVWNSANKRFOAEGGYLVPOIGDRDLDCPRARPP 64
 Db 24 GPG-----GALG-----NRNAVWNSMOTLRK--GYTVQVNVNDIYICPHYN 67
 QY 65 ARPPGPHSP-----NYEYKLYLVGAQGRRCCEAPPAPNLLTCRPRDL-----DLRTIKF 117
 Db 68 SSGVGAGAGPFGGAGABQVLYIYVSRNGYRVCNAGQGR--RWEQCRPHAPSPPIKSEKF 126

QY 118 QGVSNLMGHEFRSHHDYIIATSDGTREGLESLQGVCLTRGKVLRLRVGSPRGAAVP 177
 DB 127 QRYAFSLGVEFFHAGHEYYITSTPHNIH-----WKCLR--WKVFCVCASTSHSG--- 174
 QY 178 RKPVSSEMP-----MERDRGAHSLR-----PKENILP 204
 DB 175 EKPVPITLQFTMGPNVKINVLDEDEGENPQVPLKEKISIGSTSPFRKHLTP 223

RESULT 8

A54984
 ERF-1 protein precursor - mouse
 N:Alternate names: Cdk7 ligand
 C:Species: Mus musculus (house mouse)
 C>Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 29-Sep-1995
 C:Accession: A54984; A58873
 R:Cheng, H.J.; Flanagan, J.G.
 Cell 79, 157-168, 1994
 A:Title: Identification and cloning of ERF-1, a developmentally expressed ligand for the
 A:Reference number: A54984; PMID:9500776; PMID:7522971
 A:Accession: A54984
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-209 <CHE>
 A:Cross-references: GB:U14941; NID:G558836; PIDN:AAA53636.1; PID:G558837
 J:Shao, H.; Lou, L.; Pandey, A.; Verdierame, M.F.; Stever, D.A.; Dixit, V.M.
 J: Biol. Chem. 270, 3467-3470, 1995
 A:Title: cDNA cloning and characterization of a Cdk7 receptor protein-tyrosine kinase 14
 A:Reference number: A58873; PMID:95181289; PMID:7876076
 A:Accession: A58873
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-209 <SHA>
 A:Cross-references: GB:U14752; NID:G681886; PIDN:AAA68520.1; PID:G681887
 C:Superfamily: axon guidance signal protein
 C:Keywords: lipoprotein, membrane protein

Query Match 9.5%; Score 176; DB 2; Length 209;
 Best Local Similarity 29.3%; Pred. No. 2e-05;
 Matches 58; Conservative 19; Mismatches 69; Indels 52; Gaps 7;

QY 33 VYVNSANKRFQAE-----GGVYLPIQIGDRLLDLCPRARPSPHSSPNVEFYKLYVGGA 87
 DB 35 VYVNSNSNRFQVSAVGDGGYTVESINDYLDIYCPHGALP-PAEMMEHYLLVMVNGE 93
 QY 88 QGRCEAPPAPNLLITCDPDL---DLRFTIKFQEVSPNLMGHEFRSHHDYIIATSDGT 144
 DB 94 GHASCDHQRQGRKWEKRCRPAAPGGLKFSKFLFTFPLSGFEFRPGHEYYISATP-- 151
 QY 145 REGLESLOGVCLTRGKVLRLRVGSPRGAAVPRKPVSEMPMERDRGAHSLFEGKENLP 204
 DB 152 ---FNLYDRPCLR--LKVVYR-----PTNETLY 174
 QY 205 GDP---TSNATSRGAEG 218
 DB 175 EAPPEIFTSNSCSGLIG 192

RESULT 9

J50322
 ephrin-A2 - human
 C:Species: Homo sapiens (man)
 C>Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 21-Jul-2000
 C:Accession: U50322
 R:Adnrein, H.; Pedutour, F.; Grosgeorge, J.; Logtenberg, T.
 Biochem. Biophys. Res. Commun. 252, 378-382, 1998
 A:Title: Cloning, chromosomal mapping, and tissue expression of the gene encoding the huma
 A:Reference number: J50322; PMID:99045414; PMID:9826538
 A:Accession: J50322
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-213 <AAS>
 A:Cross-references: GB:A0007292; NID:G3688367; PIDN:CAA07435.1; PID:G3688368

C:Superfamily: axon guidance signal protein

Query Match 9.5%; Score 175.5; DB 2; Length 213;
 Best Local Similarity 36.8%; Pred. No. 2.2e-05;
 Matches 43; Conservative 14; Mismatches 51; Indels 9; Gaps 3;

QY 33 VYVNSANKRFQAE-----GGVYLPIQIGDRLLDLCPRARPSPHSSPNVEFYKLYVGGA 87
 DB 39 VYVNSNSNRFQVSAVGDGGYTVESINDYLDIYCPHGALP-PAEMMEHYLLVMVNGE 97
 QY 88 QGRCEAPPAPNLLITCDPDL---DLRFTIKFQEVSPNLMGHEFRSHHDYIIATSDGT 141
 DB 98 GHASCDHQRQGRKWEKRCRPAAPGGLKFSKFLFTFPLSGFEFRPGHEYYISATP 154

RESULT 10

A57084
 regulatory axon guidance signal protein RAGS precursor - chicken
 C:Species: Gallus gallus (chicken)
 C>Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 20-Jun-2000
 C:Accession: A57084
 R:Drescher, U.; Kremsner, C.; Handwerker, C.; Loeschinger, J.; Noda, M.; Bonhoeffer, F.
 Cell 82, 359-370, 1995
 A:Title: In vitro guidance of retinal ganglion cell axons by RAGS, a 25 kDa retinal prot
 A:Reference number: A57084; PMID:95360980; PMID:7634326
 A:Accession: A57084
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-228 <DRS>
 A:Cross-references: GB:X90377; NID:G1061113; PIDN:CAA62027.1; PID:G984118
 C:Superfamily: axon guidance signal protein
 C:Keywords: glycoprotein, membrane protein, phosphatidylinositol linkage
 F:1-20/Domain: signal sequence #status predicted <SIG>

Query Match 9.2%; Score 170.5; DB 2; Length 228;
 Best Local Similarity 28.9%; Pred. No. 5.3e-05;
 Matches 73; Conservative 30; Mismatches 91; Indels 59; Gaps 13;

QY 16 LLLIGVLTG--VSGLSLP-----VYVNSANKRFQAEVGLYLPQIGDRLLDLCPR 64
 DB 6 MLLAVNALMVCVRGQEPGRKAVADRYAVVWNTSPRFQ-QGDYHIDVINCINYLIVFCPH 64
 QY 65 ARPPGHSSPNVEFYKLYVG-----GAQGRCEAPPAPNLLITCDPDLDR 112
 DB 65 YEDSVPEKXT--ERYLYWNTFDGSSCDHISKGRKRCRPHSPN-----GPKL 113
 QY 113 FTIKFQEVSPNLMGHEFRSHHDYIIATSDGTREGLESLOGVCLTRGKVLRLRVGQ 169
 DB 114 FSEKFLFTFPLSGFEFRPGHEYYISATPONGRRS-----CLK--LKVFVR--- 159
 QY 170 SPRGAAVPRKPVSEMPMERDRGAHSLFEGKENLGDPTSNATSRGAEPLPPEKMP--A 227
 DB 160 -PANSCKMTIGHDRVFDVNDVENSLEPADDTV--RESAPSRG-ENAAQTFRIPRL 214
 QY 228 VAGAGALALLL 240
 DB 215 LITLFLMLMLI 227

RESULT 11

I38850
 LERK-4 - human
 C:Species: Homo sapiens (man)
 C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-Sep-1999
 C:Accession: I38850
 R:Kozlovsky, C.U.; Marakovsky, E.; McGrew, J.T.; Vandenbos, T.; Teepe, M.; Lyman, S.D.;
 Oncogene 10, 299-306, 1995
 A:Title: Ligands for the receptor tyrosine kinases hek and elk: isolation of cDNAs enc
 A:Reference number: I38849; PMID:95140419; PMID:7838529
 A:Accession: I38850
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-201 <RES>

A:Cross-references: EMBL:U14188, NID:g642834, PIDD:AA50079.1, PTD:g642835
 C:Genetics:
 A:Gene: GDB:EP164
 A:Cross-references: GDB:438337, OMIM:601380
 A:Map position: 1q21-q22
 C:Superfamily: axon guidance signal protein

Query Match 9.2%; Score 169.5; DB 2; Length 201;
 Best Local Similarity 29.9%; Pred. No. 5.5e-05;
 Matches 66; Conservative 18; Mismatches 82; Indels 55; Gaps 10;

Qy 24 LVGSLG-EPYVNSANKRQAGGYLYPQIGRDLDCPRARPGEPSPPYEFYKY 82
 Db 20 LRGSSLRHVYVNSNPR-LEGDVAVELGNDYDVCPEYEGGPEGP-ETFAIY 76
 Qy 83 LVGAGRCRCEAP-PAVNLLTCRDLPTIKFOEYSPNMGHFRSHDYIAT 141
 Db 77 MVMKPGYSCQAEPRYKRWVCSLPFGHQFSEKQRTFFELGFEFLGERTYIAT 136
 Qy 142 DGTREGESLQGGVCLTRGKVLIRVQSPRGAVPRKPYSEMPERDGAHSLPEKE 201
 Db 137 --TPE-----SSGQCL-----RLQVVCCKEKESAPV----- 164

Qy 202 NLPGDPTSNATS--RGAEGLPPSPMPAVAGAGLALLL 240
 Db 165 --GSPGSGISGMKGSDTPSP-----LCLELL 189

RESULT 12
 LERK-7 precursor - human
 N:Alternate names: AL-1
 C:Species: Homo sapiens (man)
 C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #ext_change 29-Sep-1999
 C:Accession: 158170; G01812
 R:Wang, J.W.; Moran, P.; Valverde, J.; Shib, A.; Yuan, J.Q.; Wong, S.C.; Tsai, S.P.;
 Neuron 14, 973-981, 1995
 A:Title: Cloning of AL-1, a ligand for an Eph-related tyrosine kinase receptor involved
 A:Reference number: 158170; PMID:95267434, PMID:7748564
 A:Accession: 158170
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-228 <RES>
 A:Cross-references: GB:577167, NID:914184, PTD:914185
 R:Kozlosky, C.J.; Vandenberg, I.; Park, U.S.; Cerretti, D.P.; Carpenter, M.K.
 submitted to the EMBL Data Library, May 1995
 A:Reference number: G08477
 A:Accession: G01812
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-228 <KOZ>
 A:Cross-references: EMBL:U26403, NID:91019430, PIDD:AA860377.1, PTD:91019431
 C:Genetics:
 A:Gene: GDB:EP167, AFI, LERK7
 A:Cross-references: GDB:568757, OMIM:601535
 A:Map position: 13q33-13q33
 C:Superfamily: axon guidance signal protein

Query Match 9.1%; Score 167.5; DB 2; Length 228;
 Best Local Similarity 28.8%; Pred. No. 8.7e-05;
 Matches 65; Conservative 29; Mismatches 81; Indels 51; Gaps 11;

Qy 33 VYVNSANKRQAGGYLYPQIGRDLDCPRARPGEPSPPYEFYKYLVG----- 85
 Db 34 VYVNSANKRQAGGYLYPQIGRDLDCPRARPGEPSPPYEFYKYLVG----- 85
 Qy 86 ---GAQGRCEAPPAVNLITCDRDLPTIKFOEYSPNMGHFRSHDYIAT 140
 Db 91 DHTSKGFKWCKRPHSPN-----GLKPEKQFLTPSLFEPFGREYIAT 141
 Qy 141 S---DGTREGESLQGGVCLTRGKVLIRVQSPRGAVPRKPYSEMPERDGAHSLPE 197
 Db 142 AIPDNGRRS-----CLK-LKVFVR---PTNSCKMTIGVHVRVVDVNDKENSLE 186

Qy 198 PKENLPDPTSNATSGAEGPLPPSPMPAVAGAGLALLLGV 243
 Db 187 PADRTV---HESAPSRG-EMAAOTPRIPRL-----LAILLFLA 223

RESULT 13
 A36377
 Bel protein precursor - human
 C:Species: Homo sapiens (man)
 C>Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #ext_change 29-Sep-1999
 C:Accession: A36377
 R:Holzman, L.B.; Marks, R.M.; Dixit, V.M.
 Mol. Cell. Biol. 10, 5830-5838, 1990
 A:Title: A novel immediate-early response gene of endothelium is induced by cytokines a
 A:Reference number: A36377; PMID:91042512, PMID:223719
 A:Accession: A36377
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-205 <HOL>
 A:Cross-references: GB:M57730, GB:M37476, NID:9179320, PIDD:AAA58388.1, PTD:9179321
 C:Superfamily: axon guidance signal protein

Query Match 9.0%; Score 166; DB 2; Length 205;
 Best Local Similarity 27.5%; Pred. No. 9.8e-05;
 Matches 46; Conservative 31; Mismatches 74; Indels 16; Gaps 5;

Qy 18 LVGVLGSLG-EPYVNSANKRQAGGYLYPQIGRDLDCPRARPGEPSPPN-- 75
 Db 8 LRGSSLRHVYVNSNPR-LEGDVAVELGNDYDVCPEYEGGPEGP-ETFAIY 76
 Qy 76 YEFKLYVNSANKRQAGGYLYPQIGRDLDCPRARPGEPSPPNMGHFRSH 132
 Db 64 MEQYLYVNSANKRQAGGYLYPQIGRDLDCPRARPGEPSPPNMGHFRSH 123
 Qy 133 HDYIATSDGTREGESLQGGVCLTRGKVLIRVQSPRGAVPRK 179
 Db 124 HSYIYISPRHQHNR-----CLRLKTVSGKITHSPQHVHPGE 163

RESULT 14
 S31216
 collagen alpha 1(X) chain precursor - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #ext_change 13-Aug-1999
 C:Accession: S31216; S28807; S22215; S30127; 148299; S26397; S31830
 R:Kong, R.Y.C.; Kwan, K.M.; Lau, E.T.; Thomas, J.T.; Boot-Handford, R.P.; Grant, M.E.;
 Eur. J. Biochem. 213, 99-111, 1993
 A:Title: Intron-exon structure, alternative use of promoter and expression of the mouse
 A:Reference number: S31216; PMID:93238750; PMID:8477738
 A:Accession: S31216
 A:Molecule type: DNA
 A:Residues: 1-680 <KON>
 A:Cross-references: EMBL:Z21610, NID:949793, PIDD:CAA79736.1, PTD:949794
 R:Elima, K.; Berola, I.; Rosati, R.; Metsaeranta, M.; Garofalo, S.; Perazaelae, M.; de Cr
 Biochem. J. 289, 247-253, 1993
 A:Title: The mouse collagen X gene: complete nucleotide sequence, exon structure and ex
 A:Reference number: S28807; PMID:93143676; PMID:8424763
 A:Accession: S28807
 A:Molecule type: DNA
 A:Residues: 1-285, 'A', 287-680 <ELI>
 A:Cross-references: EMBL:X67348, NID:950480, PIDD:CAA47763.1, PTD:950481
 R:Elima, K.; Metsaeranta, M.; Kallio, J.; Perazaelae, M.; Berola, I.; Garofalo, S.; de Cr
 Biochem. Biophys. Acta 1130, 78-80, 1992
 A:Title: Specific hybridization probes for mouse alpha-2(I)X and alpha-1(X) collagen mR
 A:Reference number: S22215; PMID:92162017; PMID:1543751
 A:Accession: S22215
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 385-450, 'K', 452-627 <ELA>
 A:Cross-references: EMBL:X63013, NID:949795, PIDD:CAA44741.1, PTD:949796
 R:Appe, S.S.; Olsen, B.R.
 Matrix 13, 165-179, 1993

Qy 198 PKENLPDPTSNATSGAEGPLPPSPMPAVAGAGLALLLGV 243
 Db 187 PADRTV---HESAPSRG-EMAAOTPRIPRL-----LAILLFLA 223
 Qy 86 ---GAQGRCEAPPAVNLITCDRDLPTIKFOEYSPNMGHFRSHDYIAT 140
 Db 91 DHTSKGFKWCKRPHSPN-----GLKPEKQFLTPSLFEPFGREYIAT 141
 Qy 141 S---DGTREGESLQGGVCLTRGKVLIRVQSPRGAVPRKPYSEMPERDGAHSLPE 197
 Db 142 AIPDNGRRS-----CLK-LKVFVR---PTNSCKMTIGVHVRVVDVNDKENSLE 186

QY 203 LPDPTSNAT-----SRGAGPLPPSPMVAAGAGLALLLVAGA 245
 Db 748 PPGPPSNGAPSGPISGPKDSDGPPERKAPRPGPFGALGAG-----LTGARGL 802
 QY 246 GAGMCMERRRPAK-----SSSRHPG-----GSFRGSGSLGCGGGMGPRAEPGE 292
 Db 803 AGPPGMFGAPSGPDPGKIKGNKPPSGNGERGPDPGLGLGTAGTAPGRDGNPGS 862
 QY 293 LGILRGG-----GADPPFCFHYEKVSGDYGHPPVTVDDGPP 330
 Db 863 DGLPGRDGAAGAKDKRGENSGPAP-----GAPGHP-----GPP 896

RESULT 16

CGH03B

collagen alpha 3(IV) chain precursor, long splice form - human
 N/Alternate names: Goodpasture antigen; procollagen alpha 3(IV) chain long splice form
 C/Species: Homo sapiens (man)
 C/Date: 28-Oct-1994 #sequence, revision 03-Oct-1995 #ext_change 22-Jun-1999
 C/Accession: A54763; A43928; A44043; A45971; A39786
 R/Mariyama, M.; Leinonen, A.; Mochizuki, T.; Tryggvason, K.; Reiders, S.T.
 J. Biol. Chem. 269, 23013-23017, 1994

A/Title: Complete primary structure of the human alpha3(IV) collagen chain. Coexpression
 A/Reference number: A54763; MUID:94364994; PMID:8083201
 A/Accession: A54763
 A/Molecule type: mRNA
 A/Residues: 1-1670 <MAR>
 A/Cross-references: GB:X80031; NID:9577563; PID:9577564
 A/Experimental source: Kidney
 R/Thurner, N.; Mason, P.T.; Brown, R.; Fox, M.; Povey, S.; Rees, A.; Pusey, C.D.
 J. Clin. Invest. 89, 592-601, 1992

A/Title: Molecular cloning of the human Goodpasture antigen demonstrates it to be the alpha
 A/Reference number: A43928; MUID:92147878; PMID:1737849
 A/Accession: A43928
 A/Molecule type: mRNA
 A/Residues: 1331-1524, '1, 1526-1670 <TUR>
 A/Cross-references: GB:M81379
 A/Experimental source: Kidney
 R/Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.
 J. Biol. Chem. 267, 19780-19784, 1992
 A/Title: Exon/intron structure of the human alpha 3(IV) gene encompassing the Goodpasture
 A/Reference number: A44043; MUID:93015826; PMID:1400291

A/Accession: A44043
 A/Molecule type: DNA, mRNA
 A/Residues: 1386-1670 <QUI>
 A/Cross-references: GB:M92993; NID:9177895; PIDN:AA21610.1; PID:9177896
 A/Note: sequence extracted from NCBI backbone (NCBI:115597)
 R/Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.
 J. Biol. Chem. 269, 17358, 1994
 A/Reference number: A44738; MUID:94274734; PMID:8006044
 A/Contents: annotation; extratun; correction to intronic sequence in A44043
 R/Bernal, D.; Quinones, S.; Saus, J.
 J. Biol. Chem. 268, 12090-12094, 1993
 A/Title: The human mRNA encoding the Goodpasture antigen is alternatively spliced.
 A/Reference number: A45971; MUID:93280184; PMID:8505332
 A/Accession: A45971
 A/Molecule type: mRNA
 A/Status: nucleic acid sequence not shown
 A/Residues: 1427-1444 <BER>
 A/Note: sequence extracted from NCBI backbone (NCBI:133163); sequence incorrectly ident
 R/Moriison, K.E.; Mariyama, M.; Yang-Feng, T.L.; Reiders, S.T.
 Am. J. Hum. Genet. 49, 545-554, 1991
 A/Title: Sequence and localization of a partial cDNA encoding the human alpha3 chain of
 A/Reference number: A39786; MUID:91353570; PMID:1882840
 A/Accession: A39786
 A/Molecule type: mRNA
 A/Residues: 1453-1593, 'A', 1595-1670 <MOR>
 A/Cross-references: GB:S55790; NID:9234418; PIDN:AA19637.1; PID:9234419
 C/Comment: Prolines and lysines at the third position of the tripeptide repeating unit
 ed and subsequently O-glycosylated.
 C/Comment: In Goodpasture's syndrome, an autoimmune response develops against an epitope

C/Genetics:

A/Gene: GDB:COL4A3

A/Cross-references: GDB:128351; OMIM:120070

A/Map position: 2q36-2q37

A/Intons: 1385/1, 1418/1, 1486/1, 1547/2, 1585/3, 1643/2 #status incomplete

A/Note: the alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands wit

C/Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha 3

monomer amino-terminal domains (with disulfide and desmosine cross-links), dimeric

er associations in the interrupted helical domain (with disulfide and desmosine cross-1

C/Function:

A/Description: minor structural component of extracellular basement membrane in kidney

C/Superfamily: collagen alpha 1(IV) chain

C/Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extrac

F/1-28/Domain: signal sequence #status predicted <SIG>

F/29-1670/Product: collagen alpha 3(IV) chain, long splice form #status predicted <MAT>

F/29-42/Domain: amino-terminal nonhelical, NH1 <NH1>

F/43-1438/Region: interrupted helical

F/791-793/Region: cell attachment (R-G-D) motif

F/996-998/Region: cell attachment (R-G-D) motif

F/1154-1156/Region: cell attachment (R-G-D) motif

F/1306-1308/Region: cell attachment (R-G-D) motif

F/1345-1347/Region: cell attachment (R-G-D) motif

F/1432-1434/Region: cell attachment (R-G-D) motif

F/1439-1670/Domain: carboxyl-terminal nonhelical, NCL <NCL>

F/1451-1551/Domain: collagen IV carboxyl-terminal repeat <CT1>

F/1561-1665/Domain: collagen IV carboxyl-terminal repeat <CT2>

F/31.33.39.41.125.422.476.479.682.722.809.1387/Disulfide bonds: interchain #status pred

F/125/Binding site: carboxylate (Asn) (covalent) #status predicted

F/1460-1548.1493-1551/Disulfide bonds: (or 1460-1551, 1493-1548) #status predicted

F/1505-1511.1616-1622/Disulfide bonds: #status predicted

F/1570-1662.1604-1665/Disulfide bonds: (or 1570-1665, 1604-1662) #status predicted

Query/Match 8.3%; Score 153.5; DB 1; Length 1670;

Best Local Similarity 27.0%; Pred. No. 0.007;

Matches 108; Conservative 26; Mismatches 155; Indels 111; Gaps 25;

QY 2 GPPHSGPGS-----VRVALLTLGLVLSGLSLPYVMNSANR-----FOAEGGYLYP 52

Db 991 GPP--GPRGDLASTGNPBPGLGLPSGMKMGMP-----GSKGRGLTGFGRGRPOLP 1044

QY 53 QI-----GDRLLDLC-----RARPGRHSSPNTEFYKLVYVGAQGRCE----- 93

Db 1045 GHHGQDQKGE-----PGYSEGRPPGPPTGDPG-----LPDGMKGMGQGPFGH 1093

QY 94 -----APRPNL--LLTCDRPLDLRT--IKQESPLTMGE-----PRSHHDYII 138

Db 1094 LPPAPPEAPSPGSPGPGKPRGPGDGFYKIGLLGPPGIRGPPGPGPSSPEYKI 1153

QY 139 ATSDTRGLLESLQGVCLTRGMKVLTVVGSPRG--GAVPRKVSSEMPERDRGAHSL 196

Db 1154 RGDQG-RDGIPEPAG-----EKGETGLLAPPGPRGNPQAQAK-----GDGA----- 1196

QY 197 EFGKENLRDPTSNATSGAGPLPPSPMVAAGAGLALLLVAGAAGMCMERRRA 256

Db 1197 -GFPGLPG-----RKAAGDAGPRPTIRBPFGPGIPGAIIPQGTNRPP--PGSRG 1248

QY 251 KPSERHHPG-----GSFRGSGSLG--LQGGGMGP--REAPGELALRGG 301

Db 1249 SFGAPRPPSPSHVITGKDKGSGHGRPKGPPGTADMDPPRLGAPGTRGLP--GP 1305

QY 302 ADPPFCFHYEKVSGDYGHPPVTVDDG-----PQSPNII 336

Db 1306 RDPDGF--QGFPGVKEXKNGPFLGSGIPGPIPGKPGPV 1344

RESULT 17

B31219

collagen 2 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #ext_change 21-Jul-2000

C/Accession: B31219; T37289; T26033; T37288

R/Kramer, J.M.; Cox, G.N.; Hirth, D.
 Cell 30, 599-606, 1982

A>Title: Comparisons of the complete sequences of two collagen genes from *Caenorhabditis*
A:Reference number: A30826; MUID:83050944; PMID:7139711
A:Accession: B31219
A:Molecule type: DNA
A:Residues: 1-301 <KR2>
A:Cross-references: GB:V00148; NID:g6683; PIDN:CAA23464.1; PID:g6684
A:Accession: T37289
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-301 <KR3>
A:Cross-references: EMBL:V00148; NID:g6683; PIDN:CAA23464.1; PID:g6684
R:Colles, L.
submitted to the EMBL Data Library, December 1995
A:Reference number: Z20141
A:Accession: T26033
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-301 <WIL>
A:Cross-references: EMBL:Z68301; PIDN:CAA92620.1; GSPDB:GN00022; CESP:W01B6.7
A:Experimental source: clone W01B6
R:Kramer, J.M.; Cox, G.N.; Hirsch, D.
J. Biol. Chem. 260, 1945-1951, 1985
A>Title: Expression of the *Caenorhabditis elegans* collagen genes col-1 and col-2 is deve
A:Reference number: Z21668; MUID:85105075; PMID:2578467
A:Accession: T37288
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-301 <KR2>
A:Cross-references: EMBL:J01048; NID:g156261; PIDN:AA27990.1; PID:g156262
C:Genetics:
A:Gene: col-2
A:Map position: 4
A:introns: 94/1
C:Superfamily: unassigned collagens

Query Match 8.2%; Score 151.5; DB 2; Length 301;
Best Local Similarity 28.7%; Pred. No. 0.0015;
Matches 50; Conservative 15; Mismatches 72; Indels 37; Gaps 7;

QY 168 GSPRGAVPRKPVSEMPERDRG-----AAHSLPEKENTPGDPT--SNATSRGAE 217
DB 129 GNPAGASAPCEPVTPQPCQPGPGPPGAPGPPGPGDNGPAGSGSPGAPGP 188
QY 218 GFLPPSPNPAVAGAGGLALLLVGA---GGAMCWRRAKPSSESHRPP-GSPGRGG 273
DB 189 GAGAGAGNDGAPGAGG-----PGEPGASBOGSP-----GEPPAPGPPAPGAGNDG 236

QY 274 STGLGGGGGMPR-----EAPPELGIALRGGAADPPPCPEYKVS 316
DB 237 AGCTGPGPAGKPGPPGAGAGADGNPGGTAGKPGEGEKGLCPKCAIDG 290

RESULT 18
T32248
hypothetical protein T15B7.4 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: T32248
R:Pauley, A.; Gattung, S.
submitted to the EMBL Data Library, September 1997
A:Description: The sequence of *C. elegans* cosmid T15B7.
A:Reference number: Z21139
A:Accession: T32248
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-325 <PAU>
A:Cross-references: EMBL:AF022965; PIDN:AA69961.1; GSPDB:GN00023; CESP:T15B7.4
A:Experimental source: strain Bristol N2; clone T15B7
C:Genetics:
A:Gene: CESP:T15B7.4
A:Map position: 5
A:introns: 266/1
C:Superfamily: unassigned collagens

Query Match 8.1%; Score 149; DB 2; Length 325;
Best Local Similarity 26.0%; Pred. No. 0.0025;
Matches 92; Conservative 25; Mismatches 109; Indels 128; Gaps 23;

QY 13 VGALLLVGLVGLSGSL-----EPVWNSANKRFOAEG---VLYPOIGDRLLDLC 62
DB 1 MSASTLVTAASAAGIAIVCVFTVMINDNSFDEKIGLKKEKGEQIA--WQANI 58
QY 63 PRARPPGPHSSPVYEFYKYLIVG-----GAQRCEA-PPAPNLLTCDRDDL 111
DB 59 PTPRPSGSS-----FLGRNRQAQCNCGASRCPCPGPP 96
QY 112 RFTIKFOEYSPNLGMHFRSHHDYIATSDGRELSLQG--GYCLTRGMKVLRLVQG 169
DB 97 ----PGPGAPGQGHGPG-----LAQPSGARNINATRPQCII----- 133
QY 170 SPRGAVPRKPVSEMPERDRGAHSLPEKENTPGDPTSNATSRGAEGLPPSPNPAVA 229
DB 134 CPAGAGPAPGAP-----PGA-----PGKNNQGPAPAQSGRGP-PGPRGP--A 175
QY 230 GAAGGLALLLVAGAGAMCWRRAKPSSESHRPPGSPGRGG--STGLGGGGM----- 283
DB 176 GDHG-----SPQPGHPSPPNPGRGGRSGGLPSPSGRPPG 212
QY 284 -GREAPPELGIALRGGAADP-PFCPEYKVSQDYGHPIVYVODGPPQSPBN 335
DB 213 PGD-ACGPGPG---HSGAGSRPGPGF-----PGPSGQGHSGNDGVGAGARN 257

RESULT 19
S4286
collagen - silkworm
C:Species: *Bombyx mori* (silkworm)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C:Accession: S4286
R:Chatey, P.P.; Beeson, M.M.; Fourche, J.J.; Bosquet, G.G.
submitted to the EMBL Data Library, March 1994
A:Description: A novel collagen encoding mRNA is transiently accumulated during ecdysis
A:Reference number: S4286
A:Accession: S4286
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-569 <CHA>
A:Cross-references: EMBL:Z30348; NID:g457768; PIDN:CAA83002.1; PID:g457769
C:Superfamily: unassigned collagens

Query Match 8.1%; Score 149; DB 2; Length 569;
Best Local Similarity 29.3%; Pred. No. 0.0045;
Matches 51; Conservative 14; Mismatches 77; Indels 32; Gaps 6;

QY 171 PRGAVPRKPVSEMPERDRGAHSLPEKENTPGDPTSNATSRGAEGLPPSPNPAVAG 230
DB 1 PGGGGQPIKPGQ-----PGYEQPGQPGYPPGP-GQPGAPQPGQPGQPGQPGG 50
QY 231 AAGGLALLLVAGAGAMCWRRAKPSSESHRPP-----PGSPGRGSGLGGGGM 283
DB 51 QAG-----QPGYGGG-----QPKPAQPGHPPGQPGQPGGPGGPGGPGG 99
QY 284 GREAPPELGIALRG-----GGAADPPPCPEYKVSQDYGHPIVYVODGPPQSP 333
DB 100 PGPGQPGQGGYPGQGGQPGQPGQPGQPGQPGQPGQPGQPGQPGQPGQPGQ 153

RESULT 20
T20497
hypothetical protein F02D10.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T20497
R:Swidburne, J.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z19283

A:Accession: T20497
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-316 <WIL>
 A:Cross-references: EMBL:Z67990; PIDN:CAA91932.1; GSPDB:GNO0028; CESP:F02D10.1
 A:Experimental source: clone F02D10
 C:Genetics:
 A:Gene: CESP:F02D10.1
 A:Map position: X
 A:Introns: 56/3
 C:Superfamily: unassigned collagens

Query Match 8.0%; Score 148.5; DB 2; Length 316;
 Best Local Similarity 28.5%; Pred. No. 0.0026;
 Matches 49; Conservative 12; Mismatches 78; Indels 33; Gaps 6;

QY 168 GGGPRGAGNPRKRVSEMPERDRGAHSLPEKKNLPGD-----TSNATSGA 216
 DB 144 GASGKGAAPCEPSTPPQ-QCPAGPPPPGDPGPBPAGSPAGSPAGSPAGP 202
 QY 217 EGPLPPSPPAVAGAGLALLLVAGAGAGMCMRRRAKSESRRHPPG-GSPRGGS 275
 DB 203 PGAGAPGNDGQPGQPGG--QDGAASSAGS-----EAGPGAPPGPAGPAGP 253
 QY 276 GUGGGGCGMPR-----EAEFGELGIALRGGAADPPCTHYEKVSG 316
 DB 254 GSGSAGGPGPKPGPPAGQPGSDGNPDTAGPPGNGEGEKGIQKYCAIDG 305

RESULT 21

collagen alpha 1(IX) chain - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 25-Dec-1994 #sequence_revision 19-Apr-1996 #text_change 20-Sep-1999
 C:Accession: S42617
 R:Rokos, I.; Muragaki, Y.; Warman, M.; Olsen, B.R.
 Matrix Biol. 14, 1-8, 1994
 A:Title: Assembly and sequencing of a cDNA covering the entire mouse alpha-1(IX) collagen
 A:Reference number: S42617; WUID:9434019; PMID:8061915
 A:Accession: S42617
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-921 <ROK>
 A:Cross-references: GB:L12215; GB:L19396; NID:G192675; PIDN:AAA21834.1; PID:G192676
 C:Superfamily: unassigned collagens

Query Match 8.0%; Score 148; DB 2; Length 921;
 Best Local Similarity 24.0%; Pred. No. 0.009;
 Matches 87; Conservative 23; Mismatches 130; Indels 122; Gaps 17;

QY 2 GPPHSGPGVRGALLLVGLVGLSLSPYVNSANKRFQAGGYLYLPQIGDRDLL 61
 DB 391 GPP--GPPPS-----GTIGHDQ-----DPL 410
 QY 62 CPRAAPGPHSSPNYEFYLYLVG-----GAQGRCEAPPAENLLTCDRLDURFTK 116
 DB 411 CRRSCPPGSSGYPG-----LPGRGHGKAGKEIG----- 440
 QY 117 FQRYSNLWGEHFRSHHDYIITATSDTREGLESLOG--GVCLTGKMYLLRVGSPRG 174
 DB 441 -----PGRQGHKEEDQCELGVEVDQCPGPGQGLRGITGIVDGKEKARFDEPEPQ 495
 QY 175 AVP-----RKVSEMEMERDRGAHSLPEKKNLPGDPSNNMTS--RGAEGLPPPS 224
 DB 496 GLPGAAGDGGQPGPPEPTPEPDGRI-----QSRKIFPSPPPKDPTGLPVDGDDGPIG 550
 QY 225 MPVAVAG-----GLALT--LLGVAGAGAMCMRRRAKPE-----SRHPG--- 265
 DB 551 MPSTKEAGKPPGPDVGLAGLPVGIPGAKGVAGEKKNTPAGPKPGQLGSSGKPGQGG 610
 QY 266 -PSPFRGSLG--GGGCGMGPREAE--RGEIGIALRG--GAAPPPPHYKXVSGYV 319
 DB 611 PPGEEVPRGRGLPGSRGVPYGPBGSGPIPGKLGVSVPGLPLGPPGLGKMGKDRGVG 670

QY 320 HP 321
 DB 671 EP 672

RESULT 22

A56101
 collagen alpha 1(XVIII) chain precursor, short splice form - mouse
 N:Contains: endostatin
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Oct-1995 #sequence_revision 08-May-1998 #text_change 31-Mar-2000
 C:Accession: A56101; A58371; S72450; S65595
 R:Rehm, M.; Philjaient, T.
 J. Biol. Chem. 270, 4705-4711, 1995

A:Title: Identification of three N-terminal ends of type XVIII collagen chains and tissue homologous to rat and Drosophila fibrinogen proteins.
 A:Reference number: A56101; WUID:95181468; PMID:7876242
 A:Accession: A56101

A:Molecule type: mRNA

A:Residues: 1-103 <REH1>

A:Cross-references: GB:U11636; NID:G618427; PIDN:AAC52178.1; PID:G618428

R:Rehm, M.; Philjaient, T.
 Proc. Natl. Acad. Sci. U.S.A. 91, 4234-4238, 1994

A:Title: Alpha1(XVIII), a collagen chain with frequent interruptions in the collagenous

A:Reference number: A58371; WUID:94240112; PMID:8183994

A:Accession: A58371

A:Molecule type: mRNA

A:Residues: 1-928 <REH2>

A:Cross-references: GB:L16898; NID:G404754; PIDN:AAA7434.1; PID:G553894

R:On, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, B.R.
 submitted to the EMBL Data Library, August 1993

A:Reference number: S72450

A:Accession: S72450

A:Molecule type: mRNA

A:Residues: 26-687, 'L', 689-734, 'E', 736-751, 'R', 753-1315 <OHS>

A:Cross-references: EMBL:L22545

C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit

lated and subsequently O-glycosylated.

C:Comment: The different splice forms of collagen alpha 1(XVIII) may be involved in per-

cyteComment: Endostatin is released from collagen alpha 1(XVIII) chain by the action of u-

ay be useful in treating solid tumors.

C:Genetics:

A:Gene: MGI:Coll8a1

A:Cross-references: MGI:71175

A:Map position: 10:41.0

C:Superfamily: unassigned collagens

C:KeyWords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteogly-

F:1-25/Domain: signal sequence #status predicted <SIG>

F:24-225/Region: thrombospondin amino-terminal similarity

F:26-1315/Product: collagen alpha 1(XVIII) chain, short splice form #status predicted <

F:172-228/Disulfide bonds: #status predicted
 F:240,245,1257/Binding site: carbohydrate (Ser) (covalent) #status predicted
 F:451,454,554/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 7.9%; Score 146.5; DB 2; Length 1315;
 Best Local Similarity 23.7%; Pred. No. 0.017;
 Matches 92; Conservative 22; Mismatches 107; Indels 167; Gaps 21;

QY 2 GPPHSGPGVRYGALLLVGLG-LYSGSLSEFYVNSANKRQAGGYLVLPQIGDRDL 60
 DB 339 GPP--GPAPOQ-----GPAAPVQSPNSQPV----- 362
 QY 61 LPPRAR-PGPHSSPWFYEFYKLVGAGGRCE----- 93
 DB 363 --PGAGPFGPGPGPKD-----GTFGRDGEFGDPRDPRDTPGQPGFPGDVG 412
 QY 94 -----APPANLLITCRPDIDLFTIKFQRYSPNLMGHEPSSHHDYIAT 140
 DB 413 PGKEKGDPIGRPGPGP-----PGPPGPFRODKLFFIME 449
 QY 141 SDGTREGESLGGVCLTRGMKVLRLVQSGPRGAVPRKP--VSEMPMERDR-GAHSLE 197
 DB 450 GSGFSGDISLGG-----PRFGPGPGPGPVPCLPSPGRFGLNSYA 492
 QY 198 PCKENLPDPTSNATSRGAGEP--LP-PPSMRPAVAGAGLALLLVAGAGAMCWRER 254
 DB 493 PGPAGLPVP-----GKEGPFPGPGPGPGPKGEP-----PGVAGQGSV----- 535
 QY 255 RAKPSSRRRPGP-GSTRGSGSLGGGGGGMGPBEIPGLIALRG--GAADPFCHY 311
 DB 536 ---GVGIPGPKSGKGDLPGLGMGKSLA---GSPFVGPVGGPGPGPGPFAAGF 588
 QY 312 EKVSQDYGHYIVQ-----DGPPQSP 333
 DB 589 DMEGS-GIPLMTTARSSDGLQPGPS 615

RESULT 23

A0333
 collagen alpha 1'(II) chain precursor - African clawed frog

C/Species: Xenopus laevis (African clawed frog)

C/Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 16-Jul-1999

C/Accession: A40333

R/Su, M.W.; Suzuki, H.R.; Bleker, J.J.; Solursh, M.; Ramirez, F.

J. Cell Biol. 115, 565-575, 1991

A/Title: Expression of two nonallelic type II procollagen genes during Xenopus laevis em

A/Reference number: A40333; PMID:192011898; PMID:1918153

A/Accession: A40333

A/Status: nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 1-1492 <SUB>

A/Cross-references: GB:M63596

A/Note: This sequence is presented as substitutions relative to another sequence in a file as they replace the appropriate interpretation of the sequence figure was reconstructed

C/Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology; C/Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix

F:1253-1492/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 7.9%; Score 146.5; DB 2; Length 1492;
 Best Local Similarity 31.1%; Pred. No. 0.019;
 Matches 56; Conservative 11; Mismatches 80; Indels 33; Gaps 8;

QY 171 PRGGAVERKPVSEMMERERG-AAHSLEBKLNLPDPTSNATSGAGPLPPSMRPAVA 229
 DB 123 PRGPDPQPGSEGGPRGRGKGTGAGPRGRDGEF---GTFGNRPAGPPPGPP 178
 QY 230 GAAGLALLLG--VAGAGAMCWRER-----RAKPSRRRHPGPGSF--GRGSLGLG 278
 DB 179 GUGGNRAAMTGSGFDEKAGAGAMVWQGMGPMGRPGPGGAGPGGFGNGGPEES 238
 QY 279 GGGG-WGPP-----EAPPEGLIALRG-----GAADPFCHYKVSQDYGHP 321

DB 239 GAGGPMGRPGPPPSGKPDGDEAGKRGKSGRGPQAGARGFPTGLPGYKHGRYP 298

RESULT 24

B56101

collagen alpha 1(XVIII) chain precursor, long splice form - mouse

N/Content: collagen alpha 1(XVIII) chain precursor, medium splice form; endostatin

C/Date: 03-Oct-1995 #sequence_revision 08-May-1998 #text_change 31-Mar-2000

C/Species: Mus musculus (house mouse)

C/Accession: B56101; C56101; S72450; S65595; PNO675; A54072; A58816

R/Rehn, M.; Pihlajaniemi, T.

J. Biol. Chem. 270, 4705-4711, 1995

A/Title: Identification of three N-terminal ends of type XVIII collagen chains and their homologous to rat and Drosophila fibrinogen proteins

A/Reference number: A56101; PMID:95181468; PMID:7876242

A/Accession: B56101

A/Molecule type: mRNA

A/Residues: 1-562 <REH>

A/Cross-references: GB:U11637; NID:9618429; PID:AAC52179.1; PID:9618430

A/Experimental source: splice form clone PE17.24

A/Accession: C56101

A/Molecule type: mRNA

A/Residues: 1-239,487-562 <REH2>

A/Cross-references: GB:U11637; NID:9618429

A/Experimental source: splice form clones PE8.1, PE9, PE15.2

R/Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, B.R.

Submitted to the EMBL Data Library, August 1993

A/Reference number: S72450

A/Accession: S72450

A/Molecule type: mRNA

A/Residues: 487-1146, 'U', 1148-1193, 'F', 1195-1210, 'R', 1212-1512, 'L', 1514-1522, 'F', 1524-1

A/Cross-references: EMBL:L22445; NID:9348968; PID:AA19787.1; PID:9511298

R/Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, A.B.R.

Proc. Natl. Acad. Sci. U.S.A. 91, 4229-4233, 1994

A/Title: Isolation and sequencing of cDNAs for proteins with multiple domains of Gly-X

A/Reference number: A56370; PMID:94240111; PMID:8183893

A/Accession: S65595

A/Molecule type: mRNA

A/Residues: 487-1512, 'U', 1514-1522, 'F', 1524-1683, 'V', 1685-1774 <OH2>

A/Cross-references: EMBL:L22445

R/Rehn, M.; Muragaki, Y.; Yoshioke, H.; Inoue, H.; Minomura, Y.

Biochem. Biophys. Res. Commun. 196, 576-582, 1993

A/Title: Identification of a novel collagen chain represented by extensive interruption

A/Reference number: PNO675; PMID:94059075; PMID:8240330

A/Accession: PNO675

A/Molecule type: mRNA

A/Residues: 635-1774 <ABR>

R/Rehn, M.; Hantikka, E.; Pihlajaniemi, T.

J. Biol. Chem. 269, 13929-13935, 1994

A/Title: Primary structure of the alpha1 chain of mouse type XVIII collagen, partial st

collagen chain.

A/Reference number: A54072; PMID:94245707; PMID:8188673

A/Accession: A54072

A/Molecule type: DNA; mRNA

A/Residues: 1293-1403, 'R', 1405-1774 <REH3>

A/Cross-references: GB:U03714; NID:9487733; PID:AAA20657.1; PID:9487734

R/O'Reilly, M.S.; Boehm, T.; Shing, Y.; Fukui, N.; Vasios, G.; Lane, W.S.; Flynn, E.; F

Cell 86, 277-285, 1997

A/Title: Endostatin: an endogenous inhibitor of angiogenesis and tumor growth.

A/Reference number: A58816; PMID:97160848; PMID:9008168

A/Accession: A58816

A/Molecule type: protein

A/Residues: 1591-1610 <ORE>

A/Experimental source: hemangioendothelium cells

A/Note: Inhibits endothelial cell proliferation

C/Comment: Prolines and tyrosines at the third position of the tripeptide repeating unit

lated and subsequently O-glycosylated.

C/Comment: The different splice forms of collagen alpha 1(XVIII) may be involved in per

C/Comment: Endostatin is released from collagen alpha 1(XVIII) chain by the action of v

C/Comment: may be useful in treating solid tumors.

C/Genetic: MGI:coll1a1

A/Genetic: MGI:71175

A/Cross-references: MGI:71175

A:Map position: 10:41.0
 A:Intons: 1295/3; 1310/1; 1331/1; 1345/3; 1388/3; 1437/1; 1461/3; 1505/3; 1516/3; 1599/3
 A>Note: the list of introns is incomplete
 C:Superfamily: unassigned collagens
 C:Keywords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteoglyc
 F:1-1774/Domain: collagen alpha 1(XVII) chain precursor, long splice form #status pred
 F:1-239/487-1774/Product: collagen alpha 1(XVII) chain precursor, medium splice form #s
 F:1-24/Domain: signal sequence #status predicted <SI>
 F:161-486/Region: frizzled similarity
 F:186-812/Domain: collagenous #status predicted <CO01>
 F:823-896/Domain: collagenous #status predicted <CO02>
 F:821-104/Domain: collagenous #status predicted <CO03>
 F:1066-1148/Domain: collagenous #status predicted <CO04>
 F:1163-1204/Domain: collagenous #status predicted <CO05>
 F:1218-1290/Domain: collagenous #status predicted <CO06>
 F:1301-1333/Domain: collagenous #status predicted <CO07>
 F:1346-1369/Domain: collagenous #status predicted <CO08>
 F:1351-1383/Region: cell attachment (R-G-D) motif
 F:1377-1428/Domain: collagenous #status predicted <CO09>
 F:1442-1459/Domain: collagenous #status predicted <CO10>
 F:1591-1774/Product: endostatin #status predicted <EST>
 F:1598-1774/Region: multiplexin collagen carboxyl-terminal similarity
 F:354-361.947/Binding site: carbohydrate (asn) (covalent) #status predicted
 F:699-704.1716/Binding site: carbohydrate (Ser) (covalent) #status predicted
 F:910.913.1053/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 7.9%; Score 146.5; DB 2; Length 1774;
 Best Local Similarity 23.7%; Pred. No. 0.023;
 Matches 92; Conservative 22; Mismatches 107; Indels 167; Gaps 21;

QY 2 GPHSGGGRVAGALLLGLG-LVSGLSLEPYNNNSANKRPAEGGYLYPQIGRDL 60
 DB 798 GPP-GRAPQ-----GPAGVVOGSPNSQPV----- 821

QY 61 LCPRAR-PPGPHSSPNVEFYKLYVGAQGRCE----- 93
 DB 822 --PQAQPPPPQPPQPPGKD-----GTPGRDGPDPGDPGRPGDTGPQGGPRTGDDVG 871

QY 94 -----APPANLLTCDPDDLDRTTTFQESPLVMKHERSHHDYITAT 140
 DB 872 PKKEGDPGIGRPPPP-----PQPPPSRROKTLFFIDWE 908

QY 141 SDGTREGLESLQGVCLTRGMKYLTVGSGPRGAVPRKP--VSEMPMERDR-GAASHLE 197
 DB 909 GSGFSGGIESLRG-----PRGFPPPPPPGPGPLGPGRFGINGSTA 951

QY 198 PKXENLPDPTSNATSRGAEP--LP-PPMPAVVAGAAGLALLLGVAGAGCMWRER 254
 DB 952 PGAGAGLPGVY-----GKEGPPGFPGPPPPPKKEGP-----PGVAGKGSV----- 994

QY 255 RAKPSRRHGP-GSFRGSGSLGCGGCGMGPRAPRGELIALRG--GAADPPCPHY 311
 DB 995 ----GDVGLPDPGSGSKDLPIDMPKSGLA---GSPGVPGPPGPPGPPGPAAG 1047

QY 312 EKVSQDYGHFVYIVQ-----DGPQSP 333
 DB 1048 DMWEGS-GIPLWTTARSSDGLQGPSP 1074

RESULT 25
 S20819
 collagen alpha 3(IX) chain precursor - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 22-Nov-1993 #sequence revision 09-Mar-1996 #text change 20-Sep-1999
 C:Accession: S20819; S22429; S22918; S22238; C18856; S22241
 R:Reviewon: R.G.; Ousepenakata, M.V.; van der Reest, M.; Wayne, R.
 Eui. J. Biochem. 205, 443-449, 1992
 A>Title: Cloning of the chicken alpha-3(IX) collagen chain completes the primary structu
 A:Reference number: S20819; MUID:92241276; PMID:1572350
 A:Accession: S20819
 A:Molecule type: mRNA
 A:Residues: 1-675 <ERR>
 A:Cross-references: EMBL:X64712; NID:g63316; PIND:CAA4567.1; PID:g63317

R:Har-El, R.; Sharma, Y.D.; Aguilera, A.; Veyama, N.; Wu, J.J.; Eyre, D.R.; Juricic, L.
 J. Biol. Chem. 267, 10070-10076, 1992
 A>Title: Cloning and developmental expression of the alpha3 chain of chicken type IX co
 A:Reference number: S22429; MUID:92230586; PMID:1577778
 A:Accession: S22429
 A:Molecule type: mRNA
 A:Residues: 1-195, 'G', 197-675 <HA1>
 A:Cross-references: EMBL:M83179
 A>Note: 353-Arg, 386-Leu and 548-Arg were also found
 R:Har-El, R.; Sharma, Y.D.; Aguilera, A.; Ueyama, N.; Wu, J.J.; Eyre, D.R.; Juricic, L.
 submitted to the EMBL Data Library, February 1992
 A:Description: Cloning and developmental expression of the alpha 3 chain of chicken typ
 A:Reference number: S22918
 A:Accession: S22918
 A:Molecule type: mRNA
 A:Residues: 1-195, 'G', 197-405, 'S', 407-675 <HA2>
 A:Cross-references: EMBL:M83179; NID:9211040; PIND:AA55960.1; PID:9211041
 R:Wayne, R.; van der Reest, M.; Nimomaya, Y.; Olsen, B.R.
 Ann. N. Y. Acad. Sci. 460, 38-46, 1985
 A>Title: The structure of type IX collagen.
 A:Reference number: S22238; MUID:86185164; PMID:3868958
 A:Accession: S22238
 A:Molecule type: Protein
 A:Residues: 540-548 <MA>
 R:Nimomaya, Y.; van der Reest, M.; Wayne, R.; Lozano, G.; Olsen, B.R.
 Biochemistry 24, 4223-4229, 1985
 A>Title: Construction and characterization of cDNA encoding the alpha2 chain of chicken
 A:Reference number: A18856; MUID:86026268; PMID:2956593
 A:Accession: C18856
 A:Molecule type: Protein
 A:Residues: 540-558 <NT>
 R:Shimomaki, M.; Wright, D.W.; Irwin, M.H.; van der Reest, M.; Wayne, R.
 Ann. N. Y. Acad. Sci. 580, 1-7, 1990
 A>Title: The structure and macromolecular organization of type IX collagen in cartilage
 A:Reference number: S22241; MUID:90247791; PMID:2186687
 A:Accession: S22241
 A:Molecule type: Protein
 A:Residues: 135, 'E', 137-187, 'X', 189-191 <SH>
 A:Superfamily: unassigned collagens
 C:Keywords: coiled coil; connective tissue; disulfide bond; extracellular matrix; heter
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:122-675/Product: collagen alpha 3(IX) chain #status predicted <MA>
 F:122-24/Domain: non-collagenous NC4 #status predicted <NC4>
 F:125-161/Domain: collagenous COL3 #status predicted <CO1>
 F:162-176/Domain: non-collagenous NC3 #status predicted <NC3>
 F:177-515/Domain: collagenous COL2 #status predicted <CO2>
 F:516-546/Domain: non-collagenous NC2 #status predicted <NC2>
 F:547-657/Domain: collagenous COL1 #status predicted <CO1>
 F:658-675/Domain: non-collagenous NC1 #status predicted <NC1>
 F:122/Modified site: pyrocollidone carboxylic acid (gin) (in mature form) #status predicte
 F:137,143,146,149,152,155,179,182,185,552,558/Modified site: hydroxyproline (Pro) #stat
 F:170,174,525,658,663/Disulfide bonds: interchain #status predicted

Query Match 7.9%; Score 146; DB 2; Length 675;
 Best Local Similarity 28.8%; Pred. No. 0.0089;
 Matches 64; Conservative 9; Mismatches 71; Indels 78; Gaps 12;

QY 166 RVGSPRGANVRKPEVSEMPMERDRGAHSLERKENTLPDPTSNATSRGAAGLPPSPM 225
 DB 23 RVG--PQPPGPPPPPPSGKIDG-----EPGSGLPFP-----GKXGAPK 65

QY 226 PAVVAGAG--GLALLLGVAGAGCMWRERAKPSRRPG-----PQSPRG- 272
 DB 66 PAAAGBAGLPPLP-----GVGLTGT-----DGPDPGNPGRDGAUGPAGPPPAKGL 115

QY 273 -----GSLGCGGCGM-----GPR-----EAEFGLGIALRG----- 300

DB 116 PGPPGPPGSPSLPGNNGFRGPPGSGLPGRPPGPPGPPGGLIPEGGGDLQCPALCP 175

QY 301 -GAADPPCFPEYKVSQDYGHFVYIVQDP-----PQSPENI 336
 DB 176 PGPGRPPGMPPGFKHTGKHKEPGEIGKEGKSGRPPGPGI 217

RESULT 26

hypothetical protein F14F7.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C/Accession: T20906
R/Murray, A.
submitted to the EMBL Data Library, November 1996

A/Reference number: Z19345
A/Accession: T20906

A/Status: preliminary; translated from GB/EMBL/DBD

A/Molecule type: DNA

A/Residues: 1-305 <MW>

A/Cross-references: EMBL:Z81503, PIDD:CA04111.1, GSPDB:GN00021, CESP:F14F7.1

A/Experimental source: clone F14F7

C/Genetics:

A/Map position: 3

A/Insertions: 27/3; 49/3

C/Superfamily: unassigned collagens

Query Match 7.9%; Score 145.5; DB 2; Length 305;
Best Local Similarity 22.5%; Pred. No. 0.0041;
Matches 83; Conservative 21; Mismatches 122; Indels 143; Gaps 12;

```

QY 11 VEVGALLLG-----VLGLVSGLSLEPPVWN-----SANK----- 40
DB 5 VILKAVRLAFSAVAFVCLSVSVVCITLPMYNYVSGKRVLVQVSPCKSAKEVFTEV 64
QY 41 -----RFGAEGGVVLPQIGDRDLDCPRARPPGHSNNYFFLYLVVG 86
DB 65 NPLRASATNSTSAHAGVGY-----AQPGGGGG-----GG 97
QY 87 AGRCCEAPPAENLLTCDRLDRLFTIKFQEVSPNLMGHSFSDHYITATSDTRE 146
DB 98 GGCSTCCRCRPGAGTGPCKG-----RPGAPGAMG----- 128
QY 147 GLESLQGVCLTRGKVKVLRVQSGPRGAVPRKPYSEMPERDRGAHSLPEKLENLPD 206
DB 129 -----MPGNPKGSGSGPHFVPTPCPCPG-----RPGPPGPPGR 165
QY 207 PTSN-----ATSRGA-----EGPLPPSPMAVGAAGLALLLGVAGAGAMCW 251
DB 166 PSSDQPGPRATGGAAPRPPGPKRGARAGNSGRAGAQQPNDHAGVGGVGAAPR 225
QY 252 RRRRAKPSERSRPPGSGFGRGGSLGAGGGGN-----GREAPGSLGIALRGGAADPPF 307
DB 226 AGPRGAPGAGHGGSSGGRRPGAPGKAPGQPGRRFP-DGHPGQPGRRPGSGSGNKG 284
QY 308 CPHYEKVSG 316
DB 285 CPKYCAIDG 293

```

RESULT 27

collagen alpha 1(X) chain precursor - bovine

C/Species: Bos primigenius taurus (cattle)

C/Date: 21-Nov-1993 #sequence_revision 23-Feb-1996 #text_change 13-Aug-1999

C/Accession: S13301

R/Thomas, J.T.; Kwan, A.P.L.; Grant, M.E.; Boot-Handford, R.P.

Biochem. J. 273, 141-148, 1991

A/Title: Isolation of cDNAs encoding the complete sequence of bovine type X collagen. EV

A/Reference number: S13301; MUID:9111331; PMID:1703407

A/Accession: S13301

A/Molecule type: mRNA

A/Residues: 1-674 <THO>

A/Cross-references: EMBL:X53556; NID:9263; PIDD:CAA37624.1; PTD:9264

C/Genetics:

A/Genes: COL1A1

C/Keywords: coiled coll., extracellular matrix; glycoprotein; homotrimer

F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-674/Product: collagen alpha 1(X) chain #status predicted <MAT>
F:547-673/Domain: complement C1q carboxyl-terminal homology <CIQ>

Query Match 7.8%; Score 145; DB 2; Length 674;
Best Local Similarity 24.2%; Pred. No. 0.01;
Matches 100; Conservative 30; Mismatches 125; Indels 158; Gaps 24;

```

QY 15 ALTLGLVGLVSGLSLEPPVWNNSANKRFGAEGGVLYLPQIGDRDLDCPR----- 65
DB 6 ALTLGLVGLVSG-----VFP-----TERVQTPFG-IKGPSNKTQFFIYALIKGKVSIR 56
QY 66 -----RPPGP-----HSSPNYEFYLYLVG-----GAQR-RCEADPAPNLLIT 103
DB 57 GEGGIPGPPGAPGRRGHPSPS-----GPPGRGTGSPGQPGILGPPGPS---A 104
QY 104 CDRPDLRLFTIKFQEVSPNLMGHSFSDHYITATSDTREGSLQ----- 153
DB 105 TGKPGI-----DGLPEKQ-----GER-GLNGPKGDI GPAGLP 137
QY 154 -----GVCLTRGKVKVLRVQ-SQPRGAVPR-----KPVSEMP-----MERDRGAHSL 196
DB 138 RGPGRPPG-RPGAGISVPKPGQPTGPRGPRGPRGEGKGTSGVGLNGKGMGHCTPC 197
QY 197 EPKKNLJEDDPISNATSRGAEGLPPP-----SMPAVAGAAGLALLLGVAGAGA 248
DB 198 RPERGLPG-----PQPTGPRGPGVGVKRGKENGCLPQCPQLKDD-----OGVGERGA 245
QY 249 MCRRRAKPSERSRPP-----GPGSFGGSLGAGGGGPREAPGELGIA----- 296
DB 246 -----AGPSGQSPGPGQGBGICPKAPGIPGQPGIPGKKGPGAGTGLGAPGF 298
QY 297 -----LRG-GGADPPPCPHYEKVSGDYGHPIYIVDDP-----PQSPENI 336
DB 299 GKXPLGLKQGSGVGLPSPGAKGEGGAPGAPGAGLPSPGNNMPQGPCKI 351

```

RESULT 28

collagen alpha 1(IX) chain precursor, long splice form - human

N/Alternate names: procollagen alpha 1(IX) chain, long splice form

C/Species: Homo sapiens (man)

C/Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 20-Sep-1999

C/Accession: S13580; S23295; S02140; G35980; S21087; S74294

R/Muragaki, Y.; Kimura, T.; Nimomiya, Y.; Olsen, B.R.

Eur. J. Biochem. 192, 703-708, 1990

A/Title: The complete primary structure of two distinct forms of human alpha-1(IX) coll

A/Reference number: S13580; MUID:91006164; PMID:2209617

A/Accession: S13580

A/Molecule type: mRNA

A/Residues: 1-931 <MR>

A/Cross-references: EMBL:X54412; NID:930083; PIDD:CA38276.1; PTD:930084

R/Nimomiya, Y.; Caesteeghola, F.; Gercke, D.; Gordon, M.K.; Jacek, O.; Luvalle, P.; M

maguchi, N.; Olsen, B.R.

In Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp.79-114, Academic P

A/Title: The molecular biology of collagens with short triple-helical domains.

A/Reference number: S22243

A/Accession: S23295

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-253, V, 255-815, 835-884 <NIN>

R/Kimura, T.; Matell, M.G.; Stevens, J.W.; Goldring, M.B.; Nimomiya, Y.; Olsen, B.R.

Eur. J. Biochem. 179, 71-78, 1989

A/Title: Molecular cloning of rat and human type IX collagen cDNA and localization of t

A/Reference number: S02140; MUID:89137096; PMID:2465149

A/Accession: S02140

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 580-596, 598-620, R, 622-813, 835-884 <KIN>

R/Muragaki, Y.; Nimomiya, I.; Henney, A.; Nimomiya, Y.; Olsen, B.R.

Proc. Natl. Acad. Sci. U.S.A. 87, 2400-2404, 1990

A/Title: The alpha(IX) collagen gene gives rise to two different transcripts in both n

A/Reference number: A35980; MUID:90207204; PMID:1690886

A:Accession: C35980
 A:Molecule type: DNA
 A:Residues: 1-4:233-267 <MUR2>
 A:Cross-references: EMBL:M32135
 R:Olson, B.R.
 submitted to the EMBL Data Library, February 1990
 A:Reference number: S21087
 A:Accession: S21087
 A:Molecule type: DNA
 A:Residues: 1-4:233-248, 'T', 250-267 <OLS>
 A:Cross-references: EMBL:M32135
 R:Diab, M.; Wu, J.U.; Byre, D.R.
 Biochem. J. 314, 327-332, 1996
 A:Title: Collagen type IX from human cartilage: a structural profile of intermolecular
 A:Reference number: S64673, PMID:96195147; PMID:8660302
 A:Accession: S74294
 A:Molecule type: protein
 A:Residues: 405-417 <DIA>
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit
 ed and subsequently O-glycosylated.
 C:Genetics:
 A:Gene: GDB:COL9A1
 A:Cross-references: GDB:119794; OMIM:120210
 A:Map position: 6q12-q14
 A:Intons: 5/2; 232/3; 260/3; 267/3
 C:Function: type IX collagen may be a heterotrimer of one alpha 1(IX) chain, one alpha 2
 A:Description: structural component of extracellular fibrous polymer associated with ty
 A:Note: in chondrocytes the long splice form is predominantly produced
 C:Superfamily: unassigned collagens
 C:Keywords: alternative splicing; coiled coil; extracellular matrix; glycoprotein; heter
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-331/Product: collagen alpha 1(IX) chain, long splice form #status predicted <MAT>
 F:24-268/Domain: non-collagenous NC4 #status predicted <NC4>
 F:269-405/Domain: collagenous COL3 #status predicted <COL3>
 F:406-417/Domain: non-collagenous NC3 #status predicted <NC3>
 F:418-756/Domain: collagenous COL2 #status predicted <COL2>
 F:757-786/Domain: non-collagenous NC2 #status predicted <NC2>
 F:787-901/Product: collagenous COL1 #status predicted <COL1>
 F:902-931/Domain: non-collagenous NC1 #status predicted <NC1>
 F:171/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match 7.8%; Score 145; DB 2; Length 931;
 Best Local Similarity 22.3%; Pred. No. 0.015;
 Matches 84; Conservative 21; Mismatches 120; Indels 152; Gaps 13;

Db 391 GPP-----GTGTHDG-----DPL 410

Qy 2 GPHSGPGVGVGALLLGLVGLSLPEVYVWNSANKRFQAGGYLVLPQIGRLDL 61

Qy 62 CPAPRPPGPHSPNVEFYKLVG-----GAQGRCEAPPAENLLTCDPDLDERPTIK 116

Db 411 CPNACPPGSGVPG-----LPGRGHKXGAKGE----- 440

Qy 117 FGYSVNLWGHFRSHHDYITATSDTRGSLQGVCLTRGMKVLLRVGSGPRGAV 176

Db 441 -----PGRQGH-----KGEKGDG-----ELSEVGAQGPFGAQ 468

Qy 177 PRKPEMPERDGAHSL-EPGKENTLPDPTSNAT-----SRGAEGLPPP 223

Db 469 GLRGITGLVDEKGEKARGLDGEPFGQJPGAFGDQSGKRPGEAGPKDRKARAGRIF 528

Qy 224 SMRAVAGAGLALLLGVAGAGACWRRRAKPSSESHPG-----PSGFGSGSLG 276

Db 529 GLGPGPKDNT-----LPGVGDRDGTGMPGTGEPGKPGPPGAGLQGLPGVGLPGAAG 583

Qy 277 LGGSGGSGPRPEAGELGALRG-----GAADPPFCPHYEKVSG 316

Db 584 VAEKSGTAGPKRGQNGSKRGQGGQGPGEVGRPGQGLPGSRGLGVGSPGLPKGLG 643

Qy 317 DYGHPIYVODGPPQSP 333

Db 644 SLGSPGLPGLPGPPGLP 660

RESULT 29
 S53787
 collagen alpha chain - Paralyvinella grasselei (fragments)
 C:Species: Paralyvinella grasselei
 C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 07-May-1999
 C:Accession: S53787
 R:Gall, F.; Mann, K.; Wiedemann, H.; Engel, J.; Timpl, R.
 J. Mol. Biol. 246, 284-294, 1995
 A:Title: Structural comparison of cuticle and interstitial collagens from annelids livi
 A:Reference number: S53786; PMID:95173973; PMID:7869380
 A:Accession: S53787
 A:Molecule type: protein
 A:Residues: 1-90;91-254;255-304;305-374;375-438 <GAI>
 C:Superfamily: unassigned collagens
 C:Keywords: coiled coil; extracellular matrix; glycoprotein

Query Match 7.8%; Score 144.5; DB 2; Length 438;
 Best Local Similarity 26.5%; Pred. No. 0.0071;
 Matches 101; Conservative 15; Mismatches 124; Indels 141; Gaps 22;

Qy 2 GPP-----HSGPGVGVGALLLGLVGLVS-----GLSLPEVYVWNSANKRFQ 43

Db 105 GPPGLTSSXGPAG-PXGANGLPQTQFAGAPGXGDPGPERGLGHP--GEAGRGQ 160

Qy 44 AEGGYLVLPQ-----IGRLDLCPAPRPP-----GPHSPNVEFYKLVG 85

Db 161 RGVGALGPPPOGPERGLPGER--GLPGRGPRGARGGGERGPPSP-----G 207

Qy 86 GAQG-RCAPAPPAENLLTCDPDLRPTIKFQEYSVNLWGHFRSHHDYITATSDG 143

Db 208 GPQDSDGAGTPGAPRGARGPRG-----SDG 235

Qy 144 TREGLSLQ-----GYCLTRGMKVLLRVGSGPRGAVPRKPVSEMPERDGAHSLP 198

Db 236 -REKPPSGSXKPGKDG-----RPDGPGRG-----LPBGRGAGS--N 269

Qy 199 GKENTLPDPTSNATSRGAEGLPSPNPAVAGAAG-GLALLLGVAGAGACWRRRA 256

Db 270 GEDQNGKP-----GKGEVGPPIQGLPGLPGEAGKGLQGLGERGK--PGERG 319

Qy 257 KPSSRRPG-----PGRFRGSLGLGGGGMGRPEAPGELGALRGGAADPPFCPHYE 312

Db 320 KDGSGRHGTAGLPGPGPGV--KGERGDHGR-GEPSGPG--SKXGRGPGGXEPQ 372

Qy 313 KYSGDYGHPIYVODGPPQSP 333

Db 373 GLRGDRNTGREGEPGAPGP 393

RESULT 30
 S28774
 collagen alpha chain - tube worm (Riftia pachyptila) (fragment)
 C:Species: Riftia pachyptila
 C:Date: 22-Nov-1993 #sequence_revision 09-Mar-1996 #text_change 23-May-1997
 C:Accession: S28774; S28915; S17581
 R:Mann, K.; Gall, F.; Timpl, R.
 Eur. J. Biochem. 210, 839-847, 1992
 A:Title: Amino acid sequence and cell-adhesion activity of a fibril-forming collagen fr
 A:Reference number: S28774; PMID:9310909; PMID:1483468
 A:Accession: S28774
 A:Molecule type: protein
 A:Residues: 1-95; 'X', 97-107; 'X', 109-191; 'X', 193-260; 'X', 262-278; 'X', 280-572; 'X', 574-611,
 A:Note: we have shown the unidentified residues as Xys forming glycosylated 5-hydroxylys
 R:Mann, K.; Gall, F.; Timpl, R.
 submitted to the Protein Sequence Database, July 1992
 A:Description: Amino acid sequence and cell adhesion activity of a fibril-forming collag
 A:Reference number: S22915
 A:Accession: S22915
 A:Molecule type: protein
 A:Residues: 1-95; 'X', 97-107; 'X', 109-191; 'X', 193-260; 'X', 262-278; 'X', 280-572; 'X', 574-611,
 A:Note: 903-proline modified to 4-hydroxyproline was also found

R.Gaill, F.; Wiedemann, H.; Mann, K.; Kuehn, K.; Timpl, R.; Engel, J.
 J. Mol. Biol. 221, 209-223, 1991
 A>Title: Molecular characterization of cuticle and interstitial collagens from worms
 A:Reference number: S17581; MUID:92015209; PMID:1920405
 A:Accession: S17581
 A:Molecule type: Protein
 A:Residues: 8-45;525-545, 'X', 547-566, 'X', 568-572, 'X', 574-611, 'X', 613-618; 'X', 811-882 <CD
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit
 h 3- and 4-hydroxylated in the X-position. Lysines are 5-hydroxylated and subsequently a
 C:Complex: homotrimer
 C:Superfamily: unassigned collagens
 C:Keywords: cell binding; coiled coil; extracellular matrix; glycoprotein; homotrimer; h
 F:1-12/Domain: amino-terminal telopeptide (fragment) <NTE>
 F:13-1023/Domain: collagenous #status experimental <COL>
 F:1024-1027/Domain: carboxyl-terminal telopeptide (fragment) <CTE>
 F:121,24,123,243,273,276,285,291,303,348,381,621,645/Modified site: 4-hydroxyproline (Pro
 F:217,39,54,72,90,93,128,150,162,165,174,177,180,207,216,219,228,237,249,255,306,312,321,
 711,714,717,723,744,759,774,783,792,816,843,849,855,861,867,886,894,915,945,954,963,966
 F:153,161,165,416,551,647,773,815,1010,1013,1016,1019/Modified site: 3-hydroxyproline (P
 F:196,108,192,261,279,573,612,657,738,765,810,927,936/Modified site: 5-hydroxylysine (Lys
 F:166,108,192,261,279,573,612,657,738,765,810,927,936/Binding site: carboxylate (Lys) (C
 F:183,342,546,567,939/Modified site: 5-hydroxylysine (Lys) #status experimental
 F:351,933/Modified site: 5-hydroxylysine (Lys) (partial) #status experimental

Query Match 7.8%; Score 144; DB 2; Length 1027;
 Best Local Similarity 22.2%; Pred. No. 0.019;
 Matches 82; Conservative 24; Mismatches 113; Indels 150; Gaps 17;

46 GGVVLYPGIGRDLDCPRARPPGPHSSPNVEFYKLYLVGAQGRRC----- 93
 4 GRRYIAQVGP-----IGRRG-PPGPPSGPGQ-----GYGLRGPPSGGXPPIGK 50
 94 -APPANLLTCDRDLRFTIKFOEYSPNLMGHEFRSHDYIATSDGRELESLO 152
 51 RPPGP-----AGIAG-----KSGDGRDGPBP 75
 153 GGVVLT-----RGMKYL-----LRVQSPRG--AVPKRPVSEM 184
 76 GGIQWPGRGAAGMGPMPKRRFRIGSGSKGQSGMVGPGPGPGRPGPIPR 135
 185 PWERDRGAHSLPEKKNLPQDPTSNATSRGAEPLPPSPMAVAGAAGLAL----- 238
 136 GQTERG-----RDGKSGLP-----LRGVGLAGPFPFGPGTSGSPGFPPTPSK 183
 239 -----LLGVAGAGMCMKRRRAKPSSESRHPG-PSFGGSGSLGCGGGM----- 283
 184 GDRGSGTIGKAGGLQPGVGLSGQPGVABNGHPMPGMDGNGBERGASGESGLPPSGFP 243
 284 GPR-----EAEPELIGLIRG-----GAADPPCPHYEKVSGDYGHPIYI 324
 244 GPRGMPGTAGSGCAQAKGDGFTGEQGRFGAPGVKSSGPP-----GDVGAPGHA 294
 325 VODGPPGP 333
 295 GEAGKRGSP 303

RESULT 31

A54121
 collagen alpha-4 chain precursor - sea urchin (strongylocentrotus purpuratus)
 N:Alternate names: collagen alpha 2(IV) chain homolog
 C:Species: Strongylocentrotus purpuratus (purple urchin)
 C:Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 13-Aug-1999
 C:Accession: A54121; S44317
 R:Expositor, J.Y.; Suzuki, H.; Geoujon, C.; Garrone, R.; Solursh, M.; Ramirez, F.
 J. Biol. Chem. 269, 13167-13171, 1994
 A>Title: Identification of a cell lineage-specific gene coding for a sea urchin alpha2(I
 A:Reference number: A54121; MUID:94230414; PMID:8175744
 A:Accession: A54121
 A:Molecule type: mRNA
 A:Residues: 1-1747 <EXP>
 A:Cross-references: EMBL:X76730; NID:9483606; PIDN:CAA54146.1; PID:9483607
 C:Genetics:

A:Gene: COL4alpha
 C:Superfamily: collagen alpha 1(IV) chain

Query Match 7.7%; Score 143; DB 2; Length 1747;
 Best Local Similarity 24.1%; Pred. No. 0.04;
 Matches 84; Conservative 24; Mismatches 159; Indels 82; Gaps 14;

7 GPGVAVGALLGLVGLVSGLSLEPVYNSANKRFOAGGVLVLPQIDRLDLCPRAR 66
 266 GPRGMG-----MGATGEVGLDSYDYEKGLPGVSGRGPFGNGITGLGMGKER 321
 67 PPGPHSSPNVEFYKLYLVGAQGRRCAPAPNLLTCDRDLRFTIKFOEYSPNLMG 126
 322 -DPPQRPQYDVK--GPGSDYGMGDPGPAADI----- 352
 127 HERSHNDYIATSDGRELESLOGVCLRMKVLRLVGO-----SPRGAVPRKPV 182
 353 -----EELVGVGPDEGPPGPNPGRGSGALGLFEDQOPGYPGMGGRPRGPTG 405
 183 EMERDRGAHSLPEKKNLPQD-----TSNATSRGAEPL-----PPSPMAVAGA 232
 406 SQEGSDGEKSGRGIDPGFOETGERGNDGNFGRGKMGVSRGPPGMDSGRGQ 465
 233 GGLALLLVGAAGMCMKRRRAKPSSESRHPG-----PSFRGSGSLGCGGGMGP 285
 466 G-----FMQKGRGP-----PGRAGPAGPARSGNOSFGFPDPTGNTGLKMRGIGALG 516
 286 REAPEPELIGLRGGAADPPCPHYEK-VSGDYGHPIYVODGPPGP 333
 517 RDRPFGSKG-----ELGGIC--PFCPPGKIGVPDRGPF-----GDPGSP 554

RESULT 32

S23779
 collagen alpha 1(VIII) chain - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
 C:Accession: S23779
 R:Udagaki, Y.; Shiota, C.; Inoue, M.; Ooshima, A.; Olsen, B.R.; Nimomiya, Y.
 Eur. J. Biochem. 207, 895-902, 1992
 A>Title: Alpha-1(VIII)-collagen gene transcripts encode a short-chain collagen polypept
 A:Reference number: S23779; MUID:92362626; PMID:1499564
 A:Accession: S23779
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-743 <MUR>
 A:Cross-references: EMBL:X66976; NID:950493; PIDN:CAA47387.1; PID:91359953
 C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
 F:616-742/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 7.7%; Score 142.5; DB 1; Length 743;
 Best Local Similarity 21.3%; Pred. No. 0.017; Indels 209; Gaps 21;
 Matches 101; Conservative 29; Mismatches 136;

14 GALLLGLVGLVSGLSLEPVYNSANKRFOAGGV-----LYPQIDRLDLCPRARPG 69
 6 GRLQLGLIFIT-----LNSV--RLIQAGAYGKLPQIPQIPQIPQIPQIPQ 55
 70 ---PH-----SSPNVEFYLY-----L 83
 56 QQVPHPLPLGDLGSKEMPHMOYGEYPLDPQMKELPPVRMGKEVPPKKGKVPPLAS 115
 84 VGAQGRRC-----APPANLLTCDRDLRFTIKFOEYSPNLMGHEFRSHDYIAT 139
 116 LRGEQPRGPRGPRGP-----PGLPGHMPG-----LK 146
 140 TSDGT-----REGLESIG--GVCLTRGMKVLRLVSGSPRGAVPRKPVSEMER 190
 147 GKPDPQYDYGPPGMPGMPGKRGAMGPAK-----GEGPKGEIGPMGIPGPO 196
 191 G--AAISL-----EPKKNLPQDPTSNATSRGAEGLPPSPM----- 225
 197 GPRPFGGLPGIGRPGRGLPGQPGAKG-ERPRKSPGPRGLGPRKGEKGFGLGLG 255


```

QY 218 GELPPSPNPAVAGAGLALLLLGVAGAGAMCWRRRRAKPSERHPGP-----GSFGRG 273
DB 555 GDRPGRGPGULPGARG-----LTGNPVGQSPGKGLPGAGGDBGRRPSPSITIKQPG 609
QY 274 SLGLGGGGGMBREAREPELGTALRGGAADPPFCPHYEKVSGDYGHVYVQDGP-- 330
DB 610 TMLPGPAGSNGDPGKPEAG----NPGVPGQGAAPGKHGCHVGEYGP-----GPPGLR 659
QY 331 -----QSPG 334
DB 660 GERGEQGP 668

RESULT 34
S15435
collagen alpha 1(VIII) chain precursor - human
C/Species: Homo sapiens (man)
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Aug-1999
C/Accession: S15435
R/Muragaki, Y.; Mattei, M.G.; Yamaguchi, N.; Olsen, B.R.; Minomiy, Y.
Eur. J. Biochem. 197, 615-622, 1991
A/Title: The complete primary structure of the human alpha-1(VIII) chain and assignment
A/Reference number: S15435; MUID:91231001; PMID:2029894
A/Accession: S15435
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-744 <MUR>
A/Cross-references: EMBL:X57527; NID:G30081; PIDN:CAA04748.1; PID:G30082
C/Genetics:
A/Gene: GDB:COL8A1
A/Cross-references: GDB:128104; OMTM:120251
A/Map position: 3q11.1-3q13.2
C/Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
F/1-20/Domain: signal sequence #status predicted <SIG>
F/21-744/Product: collagen alpha 1(VIII) chain #status predicted <MAT>
F/21-117/Region: amino-terminal nonhelical
F/118-571/Region: interrupted helical
F/572-744/Region: carboxyl-terminal nonhelical
F/617-743/Domain: complement C1q carboxyl-terminal homology <CIQ>

Query Match 7.7%; Score 142; DB 2; Length 744;
Best Local Similarity 23.5%; Pred. No. 0.019;
Matches 73; Conservative 16; Mismatches 96; Indels 126; Gaps 14;

```

```

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C/Accession: S21369; S28808; S13745; S1863
R/Brabant, X.Y.Z.; Bardon, X.Y.Z.; Dani, C.
Submitted to the EMBL Data Library, April 1992
A/Reference number: S21369
A/Accession: S21369
A/Molecule type: mRNA
A/Residues: 1-1029 <IBR>
A/Cross-references: EMBL:X65582; NID:G49808; PIDN:CAA46541.1; PID:G49809
R/Brabant, A.; Bertrand, B.; Bardon, S.; Amri, E.Z.; Grimaldi, P.; Allaud, G.; Dani,
Biochem. J. 289, 141-147, 1993
A/Title: Cloning of alpha2 chain of type VI collagen and expression during mouse develo
A/Reference number: S28808; MUID:93143659; PMID:8380980
A/Accession: S28808
A/Molecule type: mRNA
A/Residues: 266-1029 <IB2>
A/Cross-references: EMBL:X62332; NID:G49906; PIDN:CAA44206.1; PID:G49907
R/Constantinou, C.D.; Ulmen, S.A.
Matrix 11, 1-9, 1991
A/Title: Structure of CDNA encoding the triple-helical domain of murine alpha-2(VI) c
rate oligonucleotides for generation of novel CDNA clones.
A/Reference number: S13745; MUID:91226374; PMID:1709252
A/Accession: S13745
A/Molecule type: mRNA
A/Residues: 266-267; S', 269-294; 'L', 296-600 <CON>
A/Cross-references: GB:I06343; NID:G192671; PIDN:AAA37441.1; PID:G192672
A/Note: the sequence from Fig. 3 is inconsistent with that from Fig. 4 in having 306-I;
C/Keywords: alternative splicing; cartilage; cell binding; connective tissue; disulfide
F/23-1029/Product: collagen alpha 2(VI) chain #status predicted <MAT>
F/29-265/Domain: globular #status predicted <NC2>
F/54-226/Domain: von Willebrand factor type A repeat homology <VWA1>
F/266-600/Domain: collagenous #status predicted <COL>
F/376-378/Region: cell attachment (R-G-D) motif
F/436-438/Region: cell attachment (R-G-D) motif
F/499-501/Region: cell attachment (R-G-D) motif
F/508-510/Region: cell attachment (R-G-D) motif
F/549-551/Region: cell attachment (R-G-D) motif
F/601-1029/Domain: globular #status predicted <NC1>
F/623-799/Domain: von Willebrand factor type A repeat homology <VWA2>
F/841-1012/Domain: von Willebrand factor type A repeat homology <VWA3>
F/150,337,640,907,963/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.7%; Score 142; DB 1; Length 1029;
Best Local Similarity 29.7%; Pred. No. 0.027;
Matches 90; Conservative 16; Mismatches 125; Indels 72; Gaps 20;

```

```

RESULT 35
S21369
collagen alpha 2(VI) chain precursor - mouse
C/Species: Mus musculus (house mouse)

```


RESULT 38

T32250
 hypothetical protein T15B7.3 - *Caenorhabditis elegans*
 C/Species: *Caenorhabditis elegans*
 C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
 C/Accession: T32250
 R/Pauley, A.; Gertung, S.
 Submitted to the EMBL Data Library, September 1997
 A/Description: The sequence of C. elegans cosmid T15B7.
 A/Reference number: Z21139
 A/Accession: T32250
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-319 <PAU>
 A/Cross-references: EMBL:AF022985; PIDN:AB69959.1; GSPDB:GN00023; CESP:T15B7.3
 A/Experimental source: strain Bristol N2; clone T15B7
 C/Genetics:
 A/Gene: CESP:T15B7.3
 A/Map position: 5
 A/Introns: 266/1
 C/Superfamily: unassigned collagens

Query Match 7.6%; Score 141; DB 2; Length 319;
 Best Local Similarity 25.6%; Pred. No. 0.0088;
 Matches 90; Conservative 27; Mismatches 111; Indels 124; Gaps 23;

```

Qy 13 VGALLLLGVGLVSGLSL-----EPYVNSANKRPQABG---YVLPQIGDRDLIC 62
Db 1 MSASTLVTAASAAGIAIVCVFTVMIFNDISFYDEKIGELKEKGEQIA--WQAMI 58
Qy 63 PRARPGPHSPNYEFYKLYVG-----GAQGRCEA-PPAPNLLITCDRDLRL 111
Db 59 PTRRPSGSS-----FLGRNKRQAEKNGESRGCPACPP-----GQPG-- 101
Qy 112 RTTKQEVSPNMGHFRSHHDYIATSDGTREGLESLOGVCCLTRGMKYLRLVGSF 171
Db 102 ---ARGAGLPIAGQ-----PGSGARINPARGRPFCIT-----CP 135
Qy 172 RGAVPRKPESEMPERDRGAHSLPEKKNLPDPTSNATSGAAGPPLPPSMRVAGA 231
Db 136 AGAPGPPGP-----PGA-----PGPGKNGQPGAPQS--GGSGP--PGRGP--AGD 177
Qy 232 AGGIALLLGVAGAGAMCWRARRAKPSRRHPG--PGSGRGG--SLGIGGGGG--MG 284
Db 178 AG-----SPQGRGHPGSGNGRGGRGRGRGPGASGRGPGG 214
Qy 285 PRARPGELGIALRGGAGADP-PFCPHYEKVSGDYGHPIYIVODGPPGPPN 335
Db 215 PAGA-PGQPG---RSGGAGTPGPGP---PGSGQPGHSGNDGVPGTNGN 257

```

RESULT 39

A34246
 collagen alpha 1(VIII) chain precursor - rabbit
 C/Species: *Oryctolagus cuniculus* (domestic rabbit)
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C/Accession: A34246
 R/Yamaguchi, N.; Benya, P. D.; van der Rest, M.; Nishimura, Y.
 J. Biol. Chem. 264, 16022-16029, 1989
 A/Title: The cloning and sequencing of alpha1(VIII) collagen cDNAs demonstrate that type
 A/Title: similar to those of type X collagen.
 A/Reference number: A34246; MID:89380199; PMID:2476437
 A/Accession: A34246
 A/Molecule type: mRNA
 A/Residues: 1-744 <YAM>
 A/Cross-references: GB:J05042; NTD:G164895; PIDN:AAA1204.1; PID:G164896
 C/Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
 F/1-20/Domain: signal sequence #status predicted <Sig>
 F/21-744/Product: collagen alpha 1(VIII) chain #status predicted <MAT>
 F/21-117/Region: amino-terminal nonhelical
 F/118-571/Region: interrupted helical
 F/572-744/Region: carboxyl-terminal nonhelical
 F/617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 7.6%; Score 141; DB 1; Length 744;
 Best Local Similarity 23.6%; Pred. No. 0.022;
 Matches 83; Conservative 29; Mismatches 108; Indels 132; Gaps 19;

```

Qy 67 PPGHSSPNYEFYKLYVG-----GAQG-RCEAPPAAPNLLITCDRDLRLFT 114
Db 198 PPGHGLPG-----IGKPGGLPQPGAKDRPKPPGP----- 233
Qy 115 IKQEVSPNMGHFRSHHDYIATSDGTR---EGLESLOG--GYCLTRGMKYLRLV 168
Db 224 ---PGLGPRK-----GEKFGMPGLPGLKGRPHGHPGPGVGLGCVG 272
Qy 169 Q-----SPRGAVPRKPESEMPERDRGAHSLPEKKNLPDPTSN 210
Db 273 KPQVTFPGPGPLGKPGPGEPPGQPIGVPGVQGPPLPGVKKPGQDGPQGPFGG 332
Qy 211 ATSRGAGPLPPEMPRVA-----GAAGLALLLVAGAGAMCWRARR--A 256
Db 333 KGEQGLPELPGLPGLPGVKKPGFPKPD---RGLGVGALGPRGKGPVGAPEGMG 387
Qy 257 KPSRRHPG--PGSGRGGSGLL---GGGGMGPR-----EAEPEGIALRGG--GAAD 304
Db 388 PPGPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGL 447
Qy 305 PP---FCHPHYKVSVDYGH-----PYIVODGPP-----QSPNII 336
Db 448 PPGIGLPGPIGPRKGEAGHKGLPGLPGLPGLPGLPGLPGLPGLPGLPGL 499

```

RESULT 40

S23298
 collagen alpha 1(VIII) chain - chicken
 C/Species: *Gallus gallus* (chicken)
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 R/Nishimura, Y.; Castagnola, P.; Gerecke, D.; Gordon, M.K.; Jacenko, O.; LuValle, P.; Mc
 maguchi, N.; Olsen, B.R.
 In: Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp.79-114, Academic Pr.
 A/Title: The molecular biology of collagens with short triple-helical domains.
 A/Reference number: S23298
 A/Accession: S23298
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-744 <NIN>
 C/Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
 F/617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 7.6%; Score 141; DB 1; Length 744;
 Best Local Similarity 23.6%; Pred. No. 0.022;
 Matches 83; Conservative 29; Mismatches 108; Indels 132; Gaps 19;

```

Qy 67 PPGHSSPNYEFYKLYVG-----GAQG-RCEAPPAAPNLLITCDRDLRLFT 114
Db 198 PPGHGLPG-----IGKPGGLPQPGAKDRPKPPGP----- 233
Qy 115 IKQEVSPNMGHFRSHHDYIATSDGTR---EGLESLOG--GYCLTRGMKYLRLV 168
Db 224 ---PGLGPRK-----GEKFGMPGLPGLKGRPHGHPGPGVGLGCVG 272
Qy 169 Q-----SPRGAVPRKPESEMPERDRGAHSLPEKKNLPDPTSN 210
Db 273 KPQVTFPGPGPLGKPGPGEPPGQPIGVPGVQGPPLPGVKKPGQDGPQGPFGG 332
Qy 211 ATSRGAGPLPPEMPRVA-----GAAGLALLLVAGAGAMCWRARR--A 256
Db 333 KGEQGLPELPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGL 447
Qy 257 KPSRRHPG--PGSGRGGSGLL---GGGGMGPR-----EAEPEGIALRGG--GAAD 304
Db 388 PPGPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLPGL 447
Qy 305 PP---FCHPHYKVSVDYGH-----PYIVODGPP-----QSPNII 336

```


A:Accession: S12851
 A:Molecule type: mRNA
 A:Residues: 1-911 <BRO>
 R:Mayne, R.; Brewton, R.G.; Mayne, P.M.; Baker, J.R.
 J. Biol. Chem. 268, 9381-9386, 1993
 A:Title: Isolation and characterization of the chains of type V/type XI collagen present
 A:Reference number: A4662; PMID:93252802; PMID:848632
 A:Accession: C4662
 A:Molecule type: protein
 A:Residues: 347-354/356-363/586-600/912-924/925-961/962-998/999-1024 <MAY>
 A:Experimental source: vitreous humor
 A:Note: sequence modified after extraction from NCBI backbone
 R:Mu, J.; Byre, D.R.
 J. Biol. Chem. 270, 18865-18870, 1995
 A:Title: Structural analysis of cross-linking domains in cartilage type XI collagen. Ins
 A:Reference number: A56978; PMID:95370194; PMID:7642541
 A:Accession: A56978
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 206-229 <WDA>
 R:Niyibizi, C.; Byre, D.R.
 Eur. J. Biochem. 224, 943-950, 1994
 A:Title: Structural characteristics of cross-linking sites in type V collagen of bone.
 A:Reference number: S48210; PMID:95010086; PMID:7925418
 A:Accession: S65864
 A:Molecule type: protein
 A:Residues: 'X', 273-298 <NTX>
 C:Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homology
 C:Keywords: hydroxyproline
 P:211/223/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match 7.5%; Score 139; DB 2; Length 1024;
 Best Local Similarity 24.3%; Pred. No. 0.043;

Matches 90; Conservative 19; Mismatches 118; Indels 144; Gaps 19;

DB 728 GPP--GPGGV-VQPGPTEBTG-----PIGR----- 751
 QY 2 GPHSGPGGVAVGALLGLVGLVSGLSLEPYWNSANKRFQAGGYVLYPQIGDRDL 61
 DB 752 -GHPGPPGPPGEG-----LPGAAGKAGKADPGQGVSGKDP----- 789
 QY 122 PNLGHEFHSNDYVLIATSDGREGLESLOGVCITRGKVLAVGOSPRGAVRKP 181
 DB 790 AGLRG-----PGR-GLPGAQAPGL-----KGEPPGPP 820
 QY 182 SEWPERDRGAHSL-----EPGKENVDPPTSNATSRGAEGLPPSPMPAVGAAG 234
 DB 821 GPVSGPGEKSGAGTAGPIGLPGRPPGQPPG--PAGEKAPGEKGPQGPAGHDGVGPGV- 878
 QY 235 LALLLVGAVGAGMCMRRRAKPSRSRHPG-----PGSFRG----- 272
 DB 879 ----LPGRAGPKAS-----PQEDGKGEIGEPQKSGKDEGGQGMFGQDDEAG 925
 QY 273 ----GSLGAGGGGKGPPEAPFGLALRGGAADPPFCFHYKVSADYGHVYVOD 327
 DB 926 PGPPPPIGLGIPG-PPEGEGNDVGPW----GPPGP-PGIGIPGVEGCLPGAAGD 979
 QY 328 GP--PGSPENI 336
 DB 980 GPPGPIGPPGL 990

RESULT 47
 A24450
 collagen alpha 2 (VIII) chain - bovine (fragment)
 N:Alternate names: 50K-A collagen (VIII), Desemmet's membrane
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 31-Mar-1988 #sequence _revision 05-Apr-1995 #text_change 03-May-1996
 C:Accession: S12898; A24450
 R:Mann, K.; Jander, R.; Korsching, E.; Kuehn, K.; Rauterberg, J.
 FEBS Lett. 273, 168-172, 1990

A:Title: The primary structure of a triple-helical domain of collagen type VIII from bo
 A:Reference number: S12898; PMID:91032171; PMID:2226849
 A:Accession: S12898
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-469 <MAN>
 R:Kapoor, R.; Bornstein, P.; Sage, E.H.
 Biochemistry 25, 3930-3937, 1986
 A:Title: Type VIII collagen from bovine Desemmet's membrane: structural characterizat
 A:Reference number: A90507; PMID:86296655; PMID:3527259
 A:Accession: A24450
 A:Molecule type: protein
 A:Residues: 1-24 <KAP>
 C:Superfamily: collagen alpha 1 (VIII) chain; complement C1q carboxyl-terminal homology

Query Match 7.5%; Score 138.5; DB 2; Length 469;
 Best Local Similarity 25.3%; Pred. No. 0.02;

Matches 69; Conservative 17; Mismatches 72; Indels 115; Gaps 14;

QY 147 GLSLQG---GVCTRGKVLIRVQSPRGAVP-----RPVSEMPERDRGA 193
 DB 216 GLPGQPPGAXGEPGRGPGGLI---GPTGYMGPLPGXPGRDPAVPGILGDRG-- 269
 QY 194 HSLRPGKENVDP-----PTSNAT-----SRGAEGLPPSPMPAVGAAG 233
 DB 270 ---EPGDDGPPGQPGQGLGPPGLPSAGLPGRKGVPGKERTGPGVGVGIRDDG 326
 QY 234 -GLALLLVGAVGAGMCMRRRAKPSRSRHPG-----GSPGAG 272
 DB 327 PSLGA---GKPLPG-----ERGLPGAHPGPGFGPXPGEFTGRPGGPAAGALGX 376
 QY 273 GSIGL-----GGGGMGR-----EAPRGELIRRG 300
 DB 377 GDGLPGQPLRPPSGIPLGQGPAGPIPGGLPLGXGPGLPPEPGEKGVGEVA--- 432
 QY 301 GAADPPFCFHYKVSADYGHVYVODGPPSP 333
 DB 433 GPPGPPGVPSPGLTGPPGP-----GPPGP 459

RESULT 48

A41182
 collagen alpha 1 (II) chain precursor - mouse

C:Species: Mus musculus (house mouse)
 C:Date: 28-May-1992 #sequence _revision 28-May-1992 #text_change 13-Aug-1999

C:Accession: A41182; A44885
 R:Metzger, M.; Toman, D.; de Crombrughe, B.; Vuorio, E.

J. Biol. Chem. 266, 16862-16869, 1991
 A:Title: Mouse type II collagen gene. Complete nucleotide sequence, exon structure, and

A:Reference number: A41182; PMID:91358489; PMID:1885613
 A:Accession: A41182

A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA

A:Residues: 1-1419 <MEX>
 A:Cross-references: GB:M65161

R:Cheah, K.S.; Lau, E.T.; Au, P.K.; Tam, P.P.
 Development 111, 945-953, 1991

A:Title: Expression of the mouse alpha 1 (II) collagen gene is not restricted to cartil
 A:Reference number: A44885; PMID:91347939; PMID:1879363

A:Accession: A44885
 A:Molecule type: DNA

A:Residues: 1-28 <CHE>
 A:Cross-references: GB:S63190; NID:9234368; PIDN:AAB19627.1; PID:9234369

A:Note: sequence extracted from NCBI backbone (NCBIN:63190, NCBI:P:63192)
 C:Superfamily: collagen alpha 1 (I) chain; fibrillar collagen carboxyl-terminal homology

C:Keywords: alternative splicing; coiled coil; extracellular matrix; glycoprotein; trit
 P:1191-1419/domain: fibrillar collagen carboxyl-terminal homology <FOC>

Query Match 7.5%; Score 138.5; DB 2; Length 1419;
 Best Local Similarity 30.1%; Pred. No. 0.066;

Matches 58; Conservative 13; Mismatches 75; Indels 47; Gaps 9;

QY 173 GGAVPKPVSEWPERDRGAHSLRPGKENVL-----PGD--PTSNATSRGAEGLPPSP 224

```

Db 430 GGAAPFPPGGERGAPGRNGF-----PGQDGLAGPKGAPGSRGSGLAGPPGAGNDGPRPQ 464
225 MPVAVAGAGGLATLLLVAGAGGAMCWRRRRAPSRRHPGP-----GSPFRGSGSLGLGCG 280
485 EPLGPGARG-----LTCRPGDAGPQGVPSGAPGEDGRRGPPGQARGQCPVMFPPGP 539
QY 281 GGMGPRRAEFGELGIA-----LRG-----GGAADPPFCPHYKVSQDYGHPIYIVQDGP 329
Db 540 KGANGEPGKKGKGLAAGPLRGLPGKDGRTGAAGPPGSPGAPGERGEGQAP-----GP 593
QY 330 -----PQSP 334
Db 594 SGFGGLGPPGP 606

```

```

RESULT 49
B41182
Collagen alpha 1(II) chain precursor (long splice form) - mouse
C/Species: Mus musculus (house mouse)
C/Date: 28-May-1992 #sequence_revision 28-May-1992 #ext_change 16-Jul-1993
C/Accession: B41182
R/Metaseq: M.; Toman, D.; de Crombrughe, B.; Vuorio, E.
J. Biol. Chem. 266, 16862-16869, 1991
A/Title: Mouse type II collagen gene. Complete nucleotide sequence, exon structure, and
A/Reference number: A41182; PMID:91358489; PMID:1885613
A/Accession: B41182
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-1487 <MET>
A/Cross-references: GB:M5161
C/Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
C/Keywords: alternative splicing; coll; coll; extracellular matrix; glycoprotein; trime
F:33-91/Domain: von Willebrand factor type C repeat homology <YMC>
F:1259-1487/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

```

```

Query Match 7.5%; Score 138.5; DB 2; Length 1487;
Best Local Similarity 30.1%; Pred. No. 0.069;
Matches 58; Conservative 13; Mismatches 75; Indels 47; Gaps 9;
QY 173 GGAAPFPPGGERGAPGRNGF-----PGQDGLAGPKGAPGSRGSGLAGPPGAGNDGPRPQ 224
Db 498 GGAAPFPPGGERGAPGRNGF-----PGQDGLAGPKGAPGSRGSGLAGPPGAGNDGPRPQ 552
225 MPVAVAGAGGLATLLLVAGAGGAMCWRRRRAPSRRHPGP-----GSPFRGSGSLGLGCG 280
485 EPLGPGARG-----LTCRPGDAGPQGVPSGAPGEDGRRGPPGQARGQCPVMFPPGP 607
QY 281 GGMGPRRAEFGELGIA-----LRG-----GGAADPPFCPHYKVSQDYGHPIYIVQDGP 329
Db 608 KGANGEPGKKGKGLAAGPLRGLPGKDGRTGAAGPPGSPGAPGERGEGQAP-----GP 661
QY 330 -----PQSP 334
Db 662 SGFGGLGPPGP 674

```

```

RESULT 50
CGH02E
collagen alpha 2(XI) chain precursor - human (fragment)
N/Alternate names: procollagen alpha 2(XI) chain
N/contains: proline/arginine-rich protein (PARP)
C/Species: Homo sapiens (man)
C/Date: 07-Jun-1990 #sequence_revision 03-Oct-1995 #ext_change 22-Jun-1999
C/Accession: S34790; A32645
R/Zhukova, N.I.; Brewton, R.G.; Mayne, R.
FEBS Lett. 326, 25-28, 1993
A/Title: Molecular cloning of PARP (proline/arginine-rich protein) from human cartilage
chain.
A/Reference number: S34790; PMID:93314796; PMID:8325574
A/Accession: S34790
A/Molecule type: mRNA
A/Residues: 1-663 <ZHI>

```

```

A/Cross-references: EMBL:L18987; NID:9306439; PID:AAA35498.1; PID:9306440
R.Kimura, T.; Cheah, K.S.E.; Chan, S.D.H.; Lui, V.C.H.; Mattei, M.G.; van der Rest, M.;
J. Biol. Chem. 264, 13910-13916, 1989
A/Title: The human alpha2(XI) collagen (COL11A2) chain. Molecular cloning of cDNA and g
A/Reference number: A32645; PMID:93314796; PMID:12760050
A/Accession: A32645
A/Molecule type: DNA, mRNA
A/Residues: 586-1546 <KIM>
A/Cross-references: GB:U04974; NID:9180714; PID:AAA52034.1; PID:9180715
A/Note: parts of this sequence were determined by protein sequencing
C/Comment: Prolines and lysines at the third position of the tripeptide repeating unit
ed and subsequently O-glycosylated.
C/Genetics:
A/Gene: GDB:COL11A2
A/Cross-references: GDB:119788; OMIM:120290
A/Map position: 6p21.3-6p21.3
A/Introns: 1302/3; 1320/3; 1332/3; 1350/3; 1440/1; 1477/3
A/Note: the list of introns is incomplete
C/Complex: type XI collagen may be a heterotrimer of two alpha 1(XI) chains (see PIR:CG
3(XI) chain (see PIR:CGH02E), initially linked by disulfide bonds among their carboxyl
med with desmosome cross-links made from lysine and allysine residues
C/Function:
A/Description: structural component of extracellular fibrous polymer associated with ce
A/Note: may play a role in controlling the lateral growth of collagen II fibrils
C/Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homology
C/Keywords: coll; coll; extracellular matrix; glycoprotein; hydroxylysine; hydroxyprol
F:1-354/Domain: non-collagenous (fragment) #status predicted <NC3>
F:1-187/Product: proline/arginine-rich PKP protein (fragment) #status predicted <PARP>
F:1255-305/Domain: collagenous, triple helix #status predicted <COL2>
F:306-342/Domain: non-collagenous #status predicted <NC2>
F:343-1356/Region: helical
F:429-441/Region: cell attachment (R-G-D) motif
F:447-449/Region: cell attachment (R-G-D) motif
F:1257-1259/Region: cell attachment (R-G-D) motif
F:1357-1360/Region: carboxyl-terminal nonhelical telopeptide
F:1361-1546/Domain: carboxyl-terminal propeptide (fragment) #status predicted <CTP>
F:1403-1546/Domain: fibrillar collagen carboxyl-terminal homology (fragment) #status at
F:109-163; 1511-1545/Disulfide bonds: #status predicted
F:319/Modified site: alllysine (lys) #status predicted
F:432; 1266/Modified site: 5-hydroxylysine (lys) #status predicted
F:432; 1266/Binding site: carboxylate (lys) (covalent) #status predicted
F:921; 933; 1008; 1017; 1035; 1038; 1290; 1296; 1305; 1317; 1320/Modified site: 4-hydroxyproline
F:928/Modified site: 4-hydroxyproline (pro) #status atypical
F:942; 1023; 1299/Modified site: 5-hydroxylysine (lys) #status experimental
F:942; 1023; 1299/Binding site: carboxylate (lys) (covalent) #status experimental
F:1427; 1433; 1450; 1459/Disulfide bonds: interchain #status predicted
F:1460/Binding site: carboxylate (asn) (covalent) #status predicted

```

```

Query Match 7.5%; Score 138.5; DB 1; Length 1546;
Best Local Similarity 31.7%; Pred. No. 0.072;
Matches 64; Conservative 18; Mismatches 71; Indels 49; Gaps 14;
QY 147 GLBSLOGGVCLTRGMKYLRLVQSGSPRGAVPRK--PYSEMWEMERDRGAASLEPKENTLP 204
Db 1183 GSEGRQGS---KAK-----GDPGALGAPGKTPVGP-----AGPAGKGPDPGLR 1224
QY 205 GDFPS-----NATSR-GAEGPLPPSPMPAVAGAGGLALL-----LVGAVAGAGGAMCWR 252
Db 1225 GLPGSVGQGRFPAATGAGPGRPGPLPGLRDAKAKGKGRHILGILGPG----- 1279
QY 253 RRRAPKPSRRHPG--GFRGRGSLG--CGGGGMPRAAEFGELGIA--LRGGAGADPPFCP 309
Db 1280 -EQGKGRGRLPGPQSGGQGGEMGIRGASGPISG--GPPGLPAPAPKAKKATGPGGP 1357
QY 310 HYER-VSGDYGHPIYIVQDGP 330
Db 1338 KGEKGVGPPEHP-----GPF 1353

```

Search completed: February 11, 2003, 12:06:08
Job time : 36.805 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 12:05:10 / Search time 14.5409 Seconds
(without alignments)
597.393 Million cell updates/sec

Title: US-10-021-121-4

Perfect score: 1850
Sequence: 1 MGPSPGPGVGVGALLILG.....PVTIVQDGPFGSPNITYKV 340

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 segs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database :

1: /pubmed2/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1850	100.0	340	12	US-10-021-121-4
2	1846	99.8	340	9	US-10-138-787-3
3	1841	99.5	455	12	US-10-021-121-2
4	632	34.2	346	12	US-10-021-121-9
5	629.5	34.0	333	10	US-09-754-105-2
6	629.5	34.0	333	10	US-09-978-339-2
7	629.5	34.0	333	12	US-10-021-121-10
8	622.5	33.6	333	9	US-10-138-787-4
9	608.5	32.9	345	9	US-10-138-787-5
10	498	26.9	89	9	US-10-138-787-13
11	498	26.9	92	10	US-09-862-179A-17
12	284.5	15.4	92	10	US-09-864-761-48262
13	205.5	11.1	136	10	US-09-864-761-48257
14	202.5	10.9	106	10	US-09-925-297-639
15	201	10.9	82	9	US-10-138-787-11
16	201	10.9	82	9	US-09-862-179A-15
17	200.5	10.8	82	9	US-10-138-787-12
18	200.5	10.8	82	10	US-09-862-179A-16
19	179	9.7	238	10	US-09-904-954-2

20	176	9.5	209	9	US-09-921-984-2	Sequence 2, Appl1
21	174.5	9.4	233	9	US-10-138-787-7	Sequence 7, Appl1
22	172.5	9.3	218	10	US-09-925-297-510	Sequence 510, App
23	169.5	9.2	201	10	US-09-904-954-4	Sequence 4, Appl1
24	169	9.1	201	9	US-10-138-787-8	Sequence 8, Appl1
25	168.5	9.1	209	9	US-10-138-787-6	Sequence 6, Appl1
26	167.5	9.1	228	8	US-08-578-684-4	Sequence 4, Appl1
27	166	9.0	205	9	US-10-138-787-10	Sequence 10, Appl1
28	164.5	8.9	204	9	US-10-001-054-48	Sequence 48, Appl1
29	164.5	8.9	204	9	US-10-028-072-288	Sequence 288, App
30	164.5	8.9	204	9	US-10-121-049-288	Sequence 288, App
31	164.5	8.9	204	9	US-10-123-904-288	Sequence 288, App
32	164.5	8.9	204	9	US-10-140-470-288	Sequence 288, App
33	164.5	8.9	204	9	US-10-175-746-288	Sequence 288, App
34	164.5	8.9	204	9	US-10-176-918-288	Sequence 288, App
35	164.5	8.9	204	9	US-10-176-921-288	Sequence 288, App
36	160.5	8.7	228	9	US-10-138-787-9	Sequence 9, Appl1
37	154	8.3	27	10	US-09-925-301-1225	Sequence 1225, Ap
38	141	7.6	28	10	US-09-862-179A-2	Sequence 2, Appl1
39	138	7.5	638	9	US-10-001-887-108	Sequence 108, App
40	130.5	7.1	380	10	US-09-823-240-10	Sequence 10, App
41	130.5	7.1	674	10	US-09-925-299-979	Sequence 979, App
42	129	7.0	595	9	US-09-854-133-187	Sequence 187, App
43	129	7.0	595	10	US-09-738-973-187	Sequence 187, App
44	129	7.0	1806	10	US-09-919-497-56	Sequence 56, Appl1
45	126.5	6.8	403	10	US-09-925-302-689	Sequence 689, App
46	126.5	6.8	520	9	US-09-978-295A-614	Sequence 614, App
47	126.5	6.8	520	9	US-09-978-697-614	Sequence 614, App
48	126.5	6.8	520	9	US-09-978-192A-614	Sequence 614, App
49	126.5	6.8	520	9	US-09-999-832A-614	Sequence 614, App
50	126.5	6.8	520	9	US-09-978-189-614	Sequence 614, App

ALIGNMENTS

RESULT 1
US-10-021-121-4
Sequence 4, Application US/10021121
Patient No. US2002014244A1
GENERAL INFORMATION:
APPLICANT: Caras, Ingrid W
TITLE OF INVENTION: A2-1 Neurotrophic Factor
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/021,121
FILING DATE: 06-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/635,130
FILING DATE: 19-Mar-1996
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, PhD., Timothy B.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P1001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/952-9881
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids

TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4
US-10-021-121-4

Query Match 100.0%; Score 1850; DB 12; Length 340;
Best Local Similarity 100.0%; Pred. No. 8,7e-126;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPBHSFGVAVGALLLIGVGLVSGLSLEPVYNSANKRFOAGGYLYPQIGDRIDL 60
DB 1 MGPBHSFGVAVGALLLIGVGLVSGLSLEPVYNSANKRFOAGGYLYPQIGDRIDL 60
QY 61 LCPBAPPPGHSPPNVEFYKLYLVGAQGRRCBAPAPNLLTCRPPDLRFTTKFOEY 120
DB 61 LCPBAPPPGHSPPNVEFYKLYLVGAQGRRCBAPAPNLLTCRPPDLRFTTKFOEY 120
QY 121 SPNLWGHFRSHHDYIIATSDGTREGLESLOGGVCLTRGMKYLRLVQSPRGAVPRKP 180
DB 121 SPNLWGHFRSHHDYIIATSDGTREGLESLOGGVCLTRGMKYLRLVQSPRGAVPRKP 180
QY 181 VSEMPERDRGAHSLPEKENVLPDPTSNATSRGAEGLPPSPMPAVAGAAGLALLL 240
DB 181 VSEMPERDRGAHSLPEKENVLPDPTSNATSRGAEGLPPSPMPAVAGAAGLALLL 240
QY 241 GVAAGAGACWRRRRAPKPSRRHPGSGFRGSGSLGLGGGGMGPREAPBGLIALRG 300
DB 241 GVAAGAGACWRRRRAPKPSRRHPGSGFRGSGSLGLGGGGMGPREAPBGLIALRG 300
QY 301 GAADPPCPHYEKVSGDYGHPIYIVODGPPSPNIIYKYV 340
DB 301 GAADPPCPHYEKVSGDYGHPIYIVODGPPSPNIIYKYV 340

RESULT 2

US-10-138-787-3

; Sequence 3, Application US/10138787
; Patent No. US20020172984A1
; GENERAL INFORMATION:
; APPLICANT: Holland, Sacha
; APPLICANT: Mbamalu, Geraldine
; APPLICANT: Pawson, Tony
; TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED
; TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR
; TITLE OF INVENTION: TYROSINE KINASES
; FILE REFERENCE: 11757.23USWO
; CURRENT APPLICATION NUMBER: US/10/138,787
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US/09/214,631
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: PCT/CA97/00473
; PRIOR FILING DATE: 1997-07-04
; PRIOR APPLICATION NUMBER: 60/021,272
; PRIOR FILING DATE: 1996-07-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 3
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-138-787-3

Query Match 99.8%; Score 1846; DB 9; Length 340;
Best Local Similarity 99.7%; Pred. No. 1,7e-125;
Matches 339; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPBHSFGVAVGALLLIGVGLVSGLSLEPVYNSANKRFOAGGYLYPQIGDRIDL 60
DB 1 MGPBHSFGVAVGALLLIGVGLVSGLSLEPVYNSANKRFOAGGYLYPQIGDRIDL 60
QY 61 LCPBAPPPGHSPPNVEFYKLYLVGAQGRRCBAPAPNLLTCRPPDLRFTTKFOEY 120
DB 61 LCPBAPPPGHSPPNVEFYKLYLVGAQGRRCBAPAPNLLTCRPPDLRFTTKFOEY 120

QY 121 SPNLWGHFRSHHDYIIATSDGTREGLESLOGGVCLTRGMKYLRLVQSPRGAVPRKP 180
DB 121 SPNLWGHFRSHHDYIIATSDGTREGLESLOGGVCLTRGMKYLRLVQSPRGAVPRKP 180
QY 181 VSEMPERDRGAHSLPEKENVLPDPTSNATSRGAEGLPPSPMPAVAGAAGLALLL 240
DB 181 VSEMPERDRGAHSLPEKENVLPDPTSNATSRGAEGLPPSPMPAVAGAAGLALLL 240
QY 241 GVAAGAGACWRRRRAPKPSRRHPGSGFRGSGSLGLGGGGMGPREAPBGLIALRG 300
DB 241 GVAAGAGACWRRRRAPKPSRRHPGSGFRGSGSLGLGGGGMGPREAPBGLIALRG 300
QY 301 GAADPPCPHYEKVSGDYGHPIYIVODGPPSPNIIYKYV 340
DB 301 GAADPPCPHYEKVSGDYGHPIYIVODGPPSPNIIYKYV 340

RESULT 3

US-10-021-121-2

; Sequence 2, Application US/10021121
; Patent No. US2002014244A1
; GENERAL INFORMATION:
; APPLICANT: Caras, Ingrid W
; TITLE OF INVENTION: A2-1 Neurotrophic Factor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Minipain (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/021,121
; FILING DATE: 06-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,130
; FILING DATE: 19-Mar-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Ph.D., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 455 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-021-121-2

Query Match 99.5%; Score 1841; DB 12; Length 455;
Best Local Similarity 100.0%; Pred. No. 5,3e-125;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPBHSFGVAVGALLLIGVGLVSGLSLEPVYNSANKRFOAGGYLYPQIGDRIDL 60
DB 1 MGPBHSFGVAVGALLLIGVGLVSGLSLEPVYNSANKRFOAGGYLYPQIGDRIDL 60
QY 61 LCPBAPPPGHSPPNVEFYKLYLVGAQGRRCBAPAPNLLTCRPPDLRFTTKFOEY 120
DB 61 LCPBAPPPGHSPPNVEFYKLYLVGAQGRRCBAPAPNLLTCRPPDLRFTTKFOEY 120
QY 121 SPNLWGHFRSHHDYIIATSDGTREGLESLOGGVCLTRGMKYLRLVQSPRGAVPRKP 180

```

Db 121 SPNMGHGFHSHDYIITATSDGTREGLESLGGVCLTRGKVLRLVCGSPRGAVPRKX 180
Qy 181 VSEMPMRDGAHAHSLR-PGKENTLPDPTSNATSRGAEGLPPEPMPAVAGAAGLA----- 236
Db 181 VSEMPMRDGAHAHSLR-PGKENTLPDPTSNATSRGAEGLPPEPMPAVAGAAGLA----- 236
Qy 241 GVAAGAGAMCRRRRRAKPSRHRPGSPFRGSGSLGAGGGGMPREAEPEGLALRG 300
Db 241 GVAAGAGAMCRRRRRAKPSRHRPGSPFRGSGSLGAGGGGMPREAEPEGLALRG 300
Qy 301 GAADPPCPHYEKVSGDYGHVYIVODGPQSPNNITY 338
Db 301 GAADPPCPHYEKVSGDYGHVYIVODGPQSPNNITY 338

RESULT 4
US-10-021-121-9
; Sequence 9, Application US/10021121
; Patent No. US2002014244A1
; GENERAL INFORMATION:
; APPLICANT: Caras, Ingrid M
; TITLE OF INVENTION: A2-1 Neurotrophic Factor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPacIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/021,121
; FILING DATE: 06-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,130
; FILING DATE: 19-Mar-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Phd., Timothy E.
; REGISTRATION NUMBER: 36,700
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/952-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 9:
; LENGTH: 346 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-10-021-121-9

Query Match 34.2%; Score 632; DB 12; Length 346;
Best Local Similarity 39.5%; Pred. No. 2,1e-38;
Matches 145; Conservative 48; Mismatches 116; Indels 50; Gaps 9;

```

```

Qy 182 SEMPERDGAHAHSLR-PGKENTLPDPTSNATSRGAEGLPPEPMPAVAGAAGLA----- 236
Db 179 SRPSKEADNTYMAQAQSGSLGSDSGKRETYNQEKSP-----GASGSSSPD 231
Qy 237 -----LILLGVAAGGA-----MWRRRRAKPSRHRPGSPFRGSGSLG 277
Db 232 GFNSKVALFAAVGAGCVIFLLIIFLVLLKLRKRRTQ-----RAALSL 282
Qy 278 -----GGGMPREAEPEGLALRGGAADPPCPHYEKVSGDYGHVYIVODGPQSP 333
Db 283 STLASPKGSSGAGTEPSDIIPLR---TTNNYCPHYEKVSGDYGHVYIVODGPQSP 339
Qy 334 PNIYKXV 340
Db 340 ANIYKXV 346

RESULT 5
US-09-754-105-2
; Sequence 2, Application US/09754105
; Patent No. US20010009768A1
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas
; APPLICANT: Reddy, Pranthitha
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING CYTOKINE DESIGNATED LERK-5
; FILE REFERENCE: 28232
; CURRENT APPLICATION NUMBER: US/09/754,105
; CURRENT FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: 09/329,531
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 333
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-754-105-2

Query Match 34.0%; Score 629.5; DB 10; Length 333;
Best Local Similarity 40.9%; Pred. No. 3,1e-38;
Matches 135; Conservative 52; Mismatches 130; Indels 13; Gaps 5;

```

```

Qy 14 GALLLGLVGLVSGSLSEPVYMNANKRFOAEGGVLYIPOIGDRLDLCPPARPPGSHS 73
Db 14 GYLMWLCTAISKSLVLEPIYMNSSKFLPGGLVLYIPOIGDRLDLCPPARPPGSHS 70
Qy 74 PNYEYKLYLVGAQGRCEAPPAENLLTCDRDLRLFTIKFOEYSPNIMGHEFSHR 133
Db 71 GQYEVYKLYVMDKQADRCITKENTPLINCAPKPDQDIFKTFKQEPSPNIMGLEFQNK 130
Qy 134 DYIITATSDGTREGLESLGGVCLTRGKVLRLVGO--SPRGAVPRKPSSEMER-DR 190
Db 131 DYIITATSDGTREGLESLGGVCLTRGKVLRLVGO--SPRGAVPRKPSSEMER-DR 190
Qy 191 GAHSLRPGKENTLPDPTSNATSRGAEGLPPEPMPAVAGAAGLA-----LILLGVAAGAMC 250
Db 191 GRSSSTSPFVKNPPOSSSTDSAGSHGNNILGSEVALFAGIASGCIIFIVITLVLL 250
Qy 251 WRRRAKPSRHRPGSPFRGSGSLGAGGGGMPREAEPEGLALRGGAADPPCPH 310
Db 251 KYRRHRHGHSPQHTTSLSLTLATPKRSGNN---NGSEPSDIIIPR---TAQSVCFH 303
Qy 311 YEKVSGDYGHVYIVODGPQSPNNITYKXV 340
Db 304 YEKVSGDYGHVYIVODGPQSPNNITYKXV 333

RESULT 6
US-09-978-339-2
; Sequence 2, Application US/09978339
; Patent No. US20020103358A1
; GENERAL INFORMATION:

```

```

1  APPLICANT: Cerretti, Douglas P.
2  Invention
3  Raddy, Pranthia
4  TITLE OF INVENTION: No. US20020103358A1el Cyclokinine Designated Lerk-5
5
6  NUMBER OF SEQUENCES: 3
7  CORRESPONDENCE ADDRESSES:
8  ADDRESSEE: Immunex Corporation
9  STREET: 51 University Street
10 City: Seattle
11 STATE: Washington
12 COUNTRY: US
13 ZIP: 98101
14
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Floppy disk
17 COMPUTER: Apple Macintosh
18 OPERATING SYSTEM: Apple 7.1
19 SOFTWARE: Microsoft Word, Version 5.1a
20 CURRENT APPLICATION DATA:
21 FILING DATE: 15-Oct-2001
22 CLASSIFICATION: <Unknown>
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: 08/271,948
25 FILING DATE: <Unknown>
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Seese, Kathryn A.
28 REGISTRATION NUMBER: 32,172
29 REFERENCE/DOCKET NUMBER: 2823
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: (206) 587-0430
32 TELEFAX: (206) 233-0644
33
34 INFORMATION FOR SEQ ID NO: 2:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 333 amino acids
37 TYPE: amino acid
38 TOPOLOGY: linear
39 MOLECULE TYPE: protein
40 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
41
42 US-09-978-339-2
43
44 Query Match 34.0%; Score 629.5; DB 10; Length 333;
45 Best Local Similarity 40.9%; Pred. No. 3,1e-38;
46 Matches 135; Conservative 52; Mismatches 110; Indels 13; Gaps 5;
47
48 QY 14 GALLLGLVGLVSGGLELPEYWNSSANKRFOAEAGGVLYYPQIGRLDLCLCPAAPPGPHSS 73
49 DB 14 GVLVPLCRLTAISKSIVLEPIYWNSSNSKFLPGQGLVLYPQIGDLIDLICPKV--DSKTV 70
50
51 QY 74 PNEYFYLYLVGAQGRGCEAPPAEPLLTGCRPDLGRFLTKFOEYSPVLMGHEPSSH 133
52 DB 71 GQHYHYKYNVWDQDQDRCTIKKENTPLNCKAKPDQDIKFTIKQESFVLMGLBFOYK 130
53
54 QY 134 DYVIATSDGTREGLESLOGGVCLTRGMKVLRLVQ--SPRGAVPRKPYSEMPWER-DR 190
55 DB 131 DYVIATSDGTREGLESLOGGVCLTRGMKVLRLVQ--SPRGAVPRKPYSEMPWER-DR 190
56
57 QY 191 GAASLSLPGKEMLPDPTENATSRGAEGLPPEPSMAYAGAAGLALLLGVAGAGANC 250
58 DB 191 GRSSSTSPYKPNPGSGSTGNSKSHGNNILGSEVALFAGIASCIIFVITITLVLL 250
59
60 QY 251 WRERRAPSPSRPGGSGFCRGGSGLGCGGGMGPREAPDELIALRGGAADPFPCFH 310
61 DB 251 KYRRRRHKHSRPHQTTLLSLTLATPRSGNN-----NGSEPSDITIIPLR---TDSVFCFH 303
62
63 QY 311 YEKYSGPYGHPIYTVODGPPQSPNNITYYV 340
64 DB 304 YEKVSGDYGHPIYIVQEMPQSPANNITYYV 333
65
66 RESULT 7
67 US-10-021-121-10
68 Sequence 10, Application US/10021121
69 Patent No. US2002014244A1

```

```

1 GENERAL INFORMATION:
2 APPLICANT: Caras, Ingrid W
3 TITLE OF INVENTION: A2-1 Neurotrophic Factor
4 NUMBER OF SEQUENCES: 10
5 CORRESPONDENCE ADDRESS:
6 ADDRESSER: Genentech, Inc.
7 STREET: 1 DNA Way
8 CITY: South San Francisco
9 STATE: California
10 COUNTRY: USA
11 ZIP: 94080
12
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk
15 COMPUTER: IBM PC compatible
16 OPERATING SYSTEM: PC-DOS/MS-DOS
17 SOFTWARE: Winpatin (Genentech)
18
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/10/021,121
21 FILING DATE: 06-Dec-2001
22 CLASSIFICATION: <Unknown>
23
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: US/08/635,130
26 FILING DATE: 19-Mar-1996
27
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Torchia, PhD., Timothy E.
30 REGISTRATION NUMBER: 36,700
31 REFERENCE/DOCKET NUMBER: P1001
32 TELECOMMUNICATION INFORMATION:
33 TELEPHONE: 650/225-8674
34 TELEFAX: 650/952-9881
35
36 INFORMATION FOR SEQ ID NO: 10:
37
38 SEQUENCE CHARACTERISTICS:
39 LENGTH: 333 amino acids
40 TYPE: Amino Acid
41 TOPOLOGY: Linear
42
43 SEQUENCE DESCRIPTION: SEQ ID NO: 10:
44
45 US-10-021-121-10
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
```


Db 1 RRRRAKPESSRRPGSGSGGSLGLGGGGMGPRAEPGELGIALRGGAADPPCPHY 60
QY 312 EKVSQDYGHPYIVODGPPSPNNIYKY 340
Db 61 EKVSQDYGHPYIVODGPPSPNNIYKY 89

RESULT 11
US-09-862-179A-17
Sequence 17, Application US/09862179A
Patent No. US20020147306A1
GENERAL INFORMATION:
APPLICANT: Lin, Danny
APPLICANT: Pawsen, Anthony
TITLE OF INVENTION: PEPTIDES THAT MODULATE THE INTERACTION OF B CLASS EPHRINS
FILE REFERENCE: MTSI-P01-009
CURRENT APPLICATION NUMBER: US/09/862,179A
CURRENT FILING DATE: 2001-05-21
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn version 3.1
SEQ ID NO 17
LENGTH: 89
TYPE: PRT
ORGANISM: Homo sapiens
US-09-862-179A-17

Query Match 26.9%; Score 498; DB 10; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.9e-29;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 252 RRRRAKPESSRRPGSGSGGSLGLGGGGMGPRAEPGELGIALRGGAADPPCPHY 311
Db 1 RRRRAKPESSRRPGSGSGGSLGLGGGGMGPRAEPGELGIALRGGAADPPCPHY 60

QY 312 EKVSQDYGHPYIVODGPPSPNNIYKY 340
Db 61 EKVSQDYGHPYIVODGPPSPNNIYKY 89

RESULT 12
US-09-864-761-48262
Sequence 48262, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm:ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 48262
LENGTH: 92
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL136092.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
OTHER INFORMATION: SWISSPROT HIT: P98172, EVALUE 4.00e-51
OTHER INFORMATION: EST_HUMAN HIT: BE562822.1, EVALUE 3.00e-50
US-09-864-761-48262

Query Match 15.4%; Score 284.5; DB 10; Length 92;
Best Local Similarity 54.2%; Pred. No. 4.1e-14;
Matches 52; Conservative 15; Mismatches 24; Indels 5; Gaps 1;

QY 42 FQAEQGYVLPQIDRLDLCPRAPPGHSSPNYEFYLYVGAQGRCEAPAPNLL 101
Db 1 FLSGKGLVYPIKIGDLIDICRAEAGR-----YEYKLYVREQAAAGSTVADPNLL 55

QY 102 LTCDRPDLDRFTIKFOEYSPNLKHEFRSHDYIT 137
Db 56 VTGNRPEQDIRFTIKFOEFSFNYMGLFEGKHDDYI 91

RESULT 13
US-09-864-761-48257
Sequence 48257, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm:ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/609,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48257
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: MAP TO AL136092.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
; OTHER INFORMATION: EST_HUMAN HIT: AA377505.1, EVALU8 2.00e-26
; OTHER INFORMATION: SWISSPROT HIT: P52795, EVALU8 4.00e-40
US-09-864-761-48257

```

```

Query Match          11.1%; Score 205.5; DB 10; Length 136;
Best Local Similarity 37.7%; Pred. No. 3e-08; Indels 39; Gaps 5;
Matches 52; Conservative 13; Mismatches 34;

```

```

QY 230 GAAGGAA-----LTLGAGAGGA-----MCMRRRRAKPSRHRPQ 266
DB 11 GAGGSSGDPDPGFNFKVALFAVAGCVIFLLIIFLTVLLKARKRKHTQQ----- 65
QY 267 GSGRGSLGL-----GGGGMGPRAEPGELGALRGGAADPPCPHREKXSGYGHV 322
DB 66 ----RAAALSLSTLASPKGSGTAGTEPSDIIPLR---TTENNYCPHYEKVSGYGHV 118
QY 323 YIVQGPSPSPNITYKV 340
DB 119 YIVQEMPQSPNITYKV 136

```

```

RESULT 14
US-09-925-297-639
; Sequence 639, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 639
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-297-639

```

```

Query Match          10.9%; Score 202.5; DB 10; Length 106;
Best Local Similarity 43.6%; Pred. No. 3.7e-08; Indels 25; Gaps 4;
Matches 48; Conservative 11; Mismatches 26;
QY 235 LALLLGVAGAGAMCWRRAKPSRHRPQSGFGSLG-----GGGGMGPRAEP 230
DB 18 LTVLLKLT-----RKRRKHTQQ-----RAAALSLSTLASPKGSGTAGTEP 59
QY 291 GELGIALRGGAADPPCPHYEKVSGDYGHVYIVQDGPSPNITYKV 340
DB 60 SDIIPLR---TTENNYCPHYEKVSGDYGHVYIVQEMPQSPNITYKV 106

```

```

RESULT 15
US-10-138-787-11
; Sequence 11, Application US/10138787
; Patent No. US20020172984A1
; GENERAL INFORMATION:
; APPLICANT: Holland, Sacha
; APPLICANT: Madanlu, Geraldine
; APPLICANT: Pawsen, Tony
; TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED
; TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR
; FILE REFERENCE: 11757.23USWO
; CURRENT APPLICATION NUMBER: US/10/138,787
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US/09/214,631
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: PCT/CA97/00473
; PRIOR FILING DATE: 1997-07-04
; PRIOR APPLICATION NUMBER: 60/021,272
; PRIOR FILING DATE: 1996-07-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-138-787-11

```

```

Query Match          10.9%; Score 201; DB 9; Length 82;
Best Local Similarity 47.3%; Pred. No. 3.6e-08; Indels 16; Gaps 3;
Matches 44; Conservative 9; Mismatches 24;
QY 252 RRRRAKPSRHRPQSGFGSLG-----GGGGMGPRAEPGELGIALRGGAADPP 307
DB 2 RKRRKHTQQ-----RAAALSLSTLASPKGSGTAGTEPSDIIPLR---TTENNY 49
QY 308 CPHYEKVSGDYGHVYIVQDGPSPNITYKV 340
DB 50 CPHYEKVSGDYGHVYIVQEMPQSPNITYKV 82

```

```

RESULT 16
US-09-862-179A-15
; Sequence 15, Application US/09862179A
; Patent No. US20020147306A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Danny
; APPLICANT: Pawsen, Anthony
; TITLE OF INVENTION: PEPTIDES THAT MODULATE THE INTERACTION OF B CLASS EPHRINS
; TITLE OF INVENTION: AND PZ DOMAINS
; FILE REFERENCE: MTS1-P01-009
; CURRENT APPLICATION NUMBER: US/09/862,179A
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-179A-15

```

US-09-862-179A-15

10.9%; Score 201; DB 10; Length 82;

Query Match
Best Local Similarity 47.3%; Pred. No. 3,6e-08;
Matches 44; Conservative 9; Mismatches 24; Indels 16; Gaps 3;

QY 252 RRRKAPSESRHPPGSGFRGGSIGL---GGGGMGPREAPGELGIALRGGAADPPF 307

DB 2 RRRHRRHTQC-----RAALSLSTLASRGSGGTAGTEPSDIIPLR---TTEKNY 49

QY 308 CPHEKVSQDYGHPYIYVODGPPSPNNIYKYV 340

DB 50 CPHEKVSQDYGHPYIYVODGPPSPNNIYKYV 82

RESULT 17

US-10-138-787-12

Sequence 12, Application US/10138787
Patent No. US20020172984A1

GENERAL INFORMATION:

APPLICANT: Holland, Sacha

APPLICANT: Mhamali, Geraldine

APPLICANT: Pawson, Tony

TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED

TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR

FILE REFERENCE: 11757.23USWO

CURRENT APPLICATION NUMBER: US/10/138,787

CURRENT FILING DATE: 2002-05-03

PRIOR APPLICATION NUMBER: US/09/214,631

PRIOR FILING DATE: 1999-03-12

PRIOR APPLICATION NUMBER: PCT/CA97/00473

PRIOR FILING DATE: 1997-07-04

PRIOR APPLICATION NUMBER: 60/021,272

NUMBER OF SEQ ID NOS: 13

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO: 12

LENGTH: 82

TYPE: PRT

ORGANISM: Homo sapiens

US-10-138-787-12

Query Match 10.8%; Score 200.5; DB 9; Length 82;

Best Local Similarity 71.7%; Pred. No. 3,9e-08;

Matches 38; Conservative 4; Mismatches 8; Indels 3; Gaps 1;

QY 288 AEPGELGIALRGGAADPPCFHYEKVSGDYGHPIYVODGPPSPNNIYKYV 340

DB 33 SEPSDIIPLR---TADSVCFCHYEKVSQDYGHPYIYVODGPPSPNNIYKYV 82

RESULT 18

US-09-862-179A-16

Sequence 16, Application US/09862179A

Patent No. US20020147306A1

GENERAL INFORMATION:

APPLICANT: Lin, Danny

APPLICANT: Pawson, Anthony

TITLE OF INVENTION: PEPTIDES THAT MODULATE THE INTERACTION OF B CLASS EPHRINS

TITLE OF INVENTION: AND PDZ DOMAINS

FILE REFERENCE: MTST-P01-009

CURRENT APPLICATION NUMBER: US/09/862,179A

NUMBER OF SEQ ID NOS: 44

SOFTWARE: PatentIn version 3.1

SEQ ID NO: 16

LENGTH: 82

TYPE: PRT

ORGANISM: Homo sapiens

US-09-862-179A-16

Query Match 10.8%; Score 200.5; DB 10; Length 82;

Best Local Similarity 71.7%; Pred. No. 3,9e-08;

Matches 38; Conservative 4; Mismatches 8; Indels 3; Gaps 1;

QY 288 AEPGELGIALRGGAADPPCFHYEKVSGDYGHPIYVODGPPSPNNIYKYV 340

DB 33 SEPSDIIPLR---TADSVCFCHYEKVSQDYGHPYIYVODGPPSPNNIYKYV 82

RESULT 19

US-09-904-954-2

Sequence 2, Application US/09904954

Patent No. US20020010325A1

GENERAL INFORMATION:

APPLICANT: BECKMANN, M. P.

TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE

RECEPTOR HEK

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSER: IMMUNEX CORPORATION

STREET: 51 UNIVERSITY STREET

CITY: SEATTLE

STATE: WASHINGTON

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: Apple Macintosh

SOFTWARE: Microsoft Word for Apple, Version 5.1a

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/904,954

FILING DATE: 12-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/240,124

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/114,426

FILING DATE: 30-AUG-1993

APPLICATION NUMBER: US 08/109,745

FILING DATE: 20-AUG-1993

ATTORNEY/AGENT INFORMATION:

NAME: SEESE, KATHRYN A.

REGISTRATION NUMBER: 32,172

REFERENCE/DOCKET NUMBER: 2814-C

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 587-0430

TELEFAX: (206) 233-0644

TELEX: 756822

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 238 amino acids

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-904-954-2

Query Match 9.7%; Score 179; DB 10; Length 238;

Best Local Similarity 28.4%; Pred. No. 4,4e-06;

Matches 65; Conservative 24; Mismatches 80; Indels 60; Gaps 12;

QY 7 GPGGVRVALLLIGVIGLSLEPYVNSANKRQAEGGYLYLQIGDRDLPLP--R 64

DB 24 GPG-----GALG-----NRHAVVNSSQHLRR--GYTVQVNVVDYLDIYCPHYN 67

QY 65 ARPPGHSPP---NYEYLYLVGAQGRCEAPAPNILLTCRPLD---DLRFTIKF 117

DB 68 SSGVPGAGPFGGGAAPQYLVVNSNGYRTCNASGCFK--RMECHNPHAPHSIKSEKF 126

QY 118 QEYSPNIMGHEPSSHDIYIATSDGTREGLESLOGGVCULTRMKVLRLVGGSPRGANVP 177

DB 127 QRYSAFSLGYEFHAGHEYYIISTPTNLH-----WKCLR--MKVFVCCASTSHSG--- 174

QY 178 RKPVSMP-----MERDRGAHSLF-----PGKENLP 204
 Db 175 EKVPPTLPQFTMGPNVKINVLDEPFGENPQVPLKESISGTSKRRHLP 223

RESULT 20
 US-09-921-984-2
 Sequence 2, Application US/09921984
 Patent No. US20020156239A1
 GENERAL INFORMATION:

APPLICANT: Flanagan, John G.
 Cheng, Hwai-Jong
 TITLE OF INVENTION: Eph Receptor Ligands, and Uses Related
 There to

NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 60 State Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII(txt)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/921,984
 FILING DATE: 31-Jul-2001
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/308,814
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 209 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULAR TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-921-984-2

Query Match 9.5%; Score 176; DB 9; Length 209;
 Best Local Similarity 29.3%; Pred. No. 6.2e-06;
 Matches 58; Conservative 19; Mismatches 69; Indels 52; Gaps 7;

QY 33 VYVNSANKRFOAE-----GGVLYLPQIGRLDLCPARPDPGSHSPNVEFYLYLVGA 87
 Db 35 VYVNSRNPQVSAVGGGVYEVSIINDYLDIYCPHYGAPLP-PAERREYILVYVNGE 93
 QY 88 QGRRCAPAPNLLITCDRDPDL--DIRFTIKQEYSPNLMGHEPFSHEDYIIAASDGT 144
 Db 94 GHASCDHRQGFKKWECNRPAAPGGLKFSKQLFTFSLGFEFRGHEYYIISATP-- 151
 QY 145 REGLESLOGVCLTRGKVLRLVQSPRGAVPRKPVSEMPMERDRGAHSLFPGKENLP 204
 Db 152 -----PNLVDRPCLK--LKYVYR-----FTMETLY 174
 QY 205 GDP-----TSNATSRGAEG 218
 Db 175 EAPPIFTSNSSCSGLG 192
 RESULT 21

US-10-138-787-7
 Sequence 7, Application US/10138787
 Patent No. US20020172984A1
 GENERAL INFORMATION:

APPLICANT: Holland, Sacha
 Applicant: Madanlu, Geraldine
 Applicant: Pawsen, Tony
 TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED
 BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR
 TITLE OF INVENTION: TYROSINE KINASES
 FILE REFERENCE: 11757.23USWO

CURRENT APPLICATION NUMBER: US/10/138,787
 CURRENT FILING DATE: 2002-05-03
 PRIOR APPLICATION NUMBER: US/09/214,631
 PRIOR FILING DATE: 1999-03-12
 PRIOR APPLICATION NUMBER: PCT/CA97/00473
 PRIOR FILING DATE: 1997-07-04
 PRIOR APPLICATION NUMBER: 60/021,272
 PRIOR FILING DATE: 1996-07-05
 NUMBER OF SEQ ID NOS: 13
 SOFTWARE: PatentIn Ver. 2.0

LENGTH: 233
 SEQ ID NO 7
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-138-787-7

Query Match 9.4%; Score 174.5; DB 9; Length 233;
 Best Local Similarity 27.4%; Pred. No. 9e-06;
 Matches 60; Conservative 24; Mismatches 74; Indels 61; Gaps 11;

QY 20 GVLGIVSLSLPEYVNSANKRFOAGGVLYLPQIGRLDLCPR-----ARPPGHSFN 75
 Db 27 GALLG-----NRHAYVWSSNOHLRE-GYVQVNVNDIYDICHINSSGAGPDPGGAE 80
 QY 76 YEFYLYLVG-----GAQGR-----CEAPPAVNLITCDRDPDLRFITKQEYSPNLMG 127
 Db 81 YLVVNSANGYRTCNASGQFPRMCGNPHAPH-----SPKFSKPRYSAFSLGY 131
 QY 128 EPRSHEDYIITATSDGTREGSLSLQGGVCLTRGKVLRLVQSPRGAVPRKPVSEMP-- 185
 Db 132 EFHAGHEYYIISTPHNLH-----WKCLR--MKVFVCCASTHSG--EKVPPTLPQF 179
 QY 186 -----MERDRGAHSLF-----PGKENLP 204
 Db 180 TMGPVVKINVLDEPFGENPQVPLKESISGTSKRRHLP 218

RESULT 22
 US-09-925-297-510
 Sequence 510, Application US/09925297
 Patent No. US20020081659A1
 GENERAL INFORMATION:

APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 FILE REFERENCE: PA105
 CURRENT APPLICATION NUMBER: US/09/925,297
 CURRENT FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: PCT/US00/05989
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/124,270
 PRIOR FILING DATE: 1999-03-12
 NUMBER OF SEQ ID NOS: 928
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 510
 LENGTH: 218
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-925-297-510

Query Match 9.3%; Score 172.5; DB 10; Length 218;
 Best Local Similarity 28.3%; Pred. No. 1.2e-05;
 Matches 51; Conservative 30; Mismatches 80; Indels 19; Gaps 6;

```

QY 8 PGCVGVALLLGVLTGSLGSL---EPYVNSANKRFOAGGGVLYLPQIGRDLDCPR 64
DB 8 PQOABAMEFLMPLILGLCCSLAADRHTVFNWSSNPKFRNE-DYTIHQULNDYDILCPH 66
QY 65 AAPPGHSPN--YFFYLYLVGAQGRCEAPPAFNLLTCDRPDL---DLRTTYQOE 119
DB 67 YE---DHSVADAMQYILLYVEHEEYQLCQFQSDQVRCQNNRSAGHGEKLSERQR 123
QY 120 YSPNMGHFRSHHDYIILATSDTREGESLQSGVCLTRGMKYLRLRVGQSPRGAVPRK 179
DB 124 FTFPLTGKFKGSHGYIYSKPIHQHEDR-----CRLKATVSGKITTHSPQAHNDNQE 176

RESULT 23
US-09-904-954-4
; Sequence 4, Application US/09904954
; Patent No. US20020010325A1
; GENERAL INFORMATION:
; APPLICANT: BECKMANN, M. P.
; CERESETTI, DOUGLAS P.
; TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
; RECEPTOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TMMNEX CORPORATION
; STREET: 51 UNIVERSITY STREET
; CITY: SEATTLE
; STATE: WASHINGTON
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: Apple Macintosh
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/904,954
; FILING DATE: 12-Jul-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/240,124
; FILING DATE: <unknown>
; APPLICATION NUMBER: US 08/114,426
; FILING DATE: 30-AUG-1993
; APPLICATION NUMBER: US 08/109,745
; FILING DATE: 20-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SEASE, KATHRYN A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2814-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEK: 756822
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-904-954-4

Query Match 9.2%; Score 169.5; DB 10; Length 201;
Best Local Similarity 29.9%; Pred. No. 1.7e-05;
Matches 66; Conservative 18; Mismatches 82; Indels 55; Gaps 10;
QY 24 LVSGSL-EPYVNSANKRFOAGGGVLYLPQIGRDLDCPRAPPGHSPNTEFFYLY 82
DB 20 LRGGSSLRHVYVWSSNPRL-LRGDAVVELGNDYLDIVCPHYEGPPEGP--ETFALY 76
QY 83 LVGAQGRCEAP-PAFNLLTCDRPDLRTIKFQESYSPMLMGHFRSHHDYIILATS 141

```

```

DB 77 MYMVPYEQAGLEGPAYRWVCSLPFGHVQFSEKIQRTPTSLGFEFFPGERTYIISVP 136
QY 142 DGTREGESLQSGVCLTRGMKYLRLRVGQSPRGAVPRKVSMPMERDRGAHSLPEPGE 201
DB 137 --TPE-----SSGQCL-----RLOVSVCCERKESAHVP----- 164
QY 202 NLPGDPTSNATS--RGAEGLPPEPSMPAVAGAAGLALLL 240
DB 165 ---GSPGESGTSGMRGDTTPSP-----LCILL 189

RESULT 24
US-10-138-787-8
; Sequence 8, Application US/10138787
; Patent No. US20020172984A1
; GENERAL INFORMATION:
; APPLICANT: Holland, Sacha
; APPLICANT: Mbamalu, Geraldine
; APPLICANT: Rawson, Tony
; TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED
; TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR
; FILE REFERENCE: 11757.23USMO
; CURRENT APPLICATION NUMBER: US/10/138,787
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US/09/214,631
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: PCT/CA97/00473
; PRIOR FILING DATE: 1997-07-04
; PRIOR APPLICATION NUMBER: 60/021,272
; PRIOR FILING DATE: 1996-07-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-138-787-8

Query Match 9.1%; Score 169; DB 9; Length 201;
Best Local Similarity 29.8%; Pred. No. 1.9e-05;
Matches 70; Conservative 19; Mismatches 84; Indels 62; Gaps 11;
QY 17 LILGVY-----GLVSGSL-EPYVNSANKRFOAGGGVLYLPQIGRDLDCPRARP 68
DB 6 LILTVMAAPLGGSLRGSSLRHVYVWSSNPR-SLRGDAVVELGNDYLDIVCPHYEGP 64
QY 69 GPRSSNRYEFYLYLVGAQGRCEAP-PAFNLLTCDRPDLRTIKFQESYSPMLMGH 127
DB 65 GPPEGP--ETFALYVDMWGYESCQLEGPGRGYRWVCSLPFGHVQFSEKIQRTPTSLGF 122
QY 128 EFRSHHDYIILATSDTREGESLQSGVCLTRGMKYLRLRVGQSPRGAVPRKVSMPYE 187
DB 123 EFLPGERTYIISVP--TPE-----SSGQCL-----RLOVSVCCERKES 155
QY 188 RDRGAHSLPEPKENLPDPTSNATS--RGAEGLPPEPSMPAVAGAAGLALLL 240
DB 156 RKESAHVP-----GSPGESGTSGMRGDTTPSP-----LCILL 189

RESULT 25
US-10-138-787-6
; Sequence 6, Application US/10138787
; Patent No. US20020172984A1
; GENERAL INFORMATION:
; APPLICANT: Holland, Sacha
; APPLICANT: Mbamalu, Geraldine
; APPLICANT: Rawson, Tony
; TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED
; TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR
; FILE REFERENCE: 11757.23USMO

```


APPLICANT: Baker, Kevin
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin
APPLICANT: Hebert, Carolyn
APPLICANT: Henzel, William
APPLICANT: Kabakoff, Rhona
APPLICANT: Shelton, David
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin
APPLICANT: Wood, William
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC
FILE REFERENCE: P3034R1PCT
CURRENT APPLICATION NUMBER: US/10/001,054
CURRENT FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/038114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/090691
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/096891
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: 60/096894
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100263
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/101476
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/107783
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: 60/108649
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112420
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/115554
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/116533
PRIOR FILING DATE: 1999-01-20
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/131294
PRIOR FILING DATE: 1999-04-07
PRIOR APPLICATION NUMBER: 60/140650
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/187202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/209832
PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 60/232887
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-11-19
PRIOR APPLICATION NUMBER: 09/218517
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 09/284291
PRIOR FILING DATE: 1999-04-12
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380913
PRIOR FILING DATE: 1999-09-09
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/866034
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/882636
PRIOR FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/924419
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/927796
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/929404
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 09/941992
PRIOR FILING DATE: 2001-08-25
PRIOR APPLICATION NUMBER: 09/946374
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: PCT/US98/18824
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: PCT/US99/00106
PRIOR FILING DATE: 1999-01-05
PRIOR APPLICATION NUMBER: PCT/US99/05028
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: PCT/US99/08615
PRIOR FILING DATE: 1999-04-20
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/20111
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00376
PRIOR FILING DATE: 2000-01-06
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04341

PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-26
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085223
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13

PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

Query Match 8.9%; Score 164.5; DB 9; Length 204;
Best Local Similarity 27.7%; Pred. No. 4,1e-05;
Matches 46; Conservative 31; Mismatches 74; Indels 15; Gaps 5;

QY 18 LIGVLGIVSGLSLEPYVYNSANKRFGAEGGYLYPQIGRDLICPRARPQPHSS-PNY 76
DB 8 LIGVCSLAADRHVTFWMSNPKFRNE-DYTHVOLNDVDTICHYE---DHSADAM 63
QY 77 EFKLYLVGAGQRCRCEAPAPNLTCDRDPD---DLRFITKFOEYSNLTWGBFRSHR.133
DB 64 EQYLIVVEHEEYQLOQPOSKQDVWQCNRPASAKHGPEKLSKFFQFTFTLGKEFKEGH 123
QY 134 DYTITATSDGTREGLESLOGGVCLTRGMKYLRLVQSGPRGAVPRK 179
DB 124 SYYYISKPIHQHEDR-----CLRUKVTVSGKITHSPQAHNDPQE 162

RESULT 30
US-10-121-049-288
Sequence 288, Application US/10121049
Publication No. US2003002229A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen

```

RESULT 31
US-10-123-904-288
; Sequence 288, Application US/10123904
; Publication No. US20030022328A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Demoyers, Luc
APPLICANT: Flivartoff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tomas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zhenh
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEOTIC
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEOTIC
FILE REFERENCE: P3330R1C54
CURRENT APPLICATION NUMBER: US/10/123,904
;
; Prior Application removed - See File Wrapper or Palm

```

```

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tunas, Daniel
APPLICANT: Matanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zhenli
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P33081C160
CURRENT APPLICATION NUMBER: US/10/140,470
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 288
LENGTH: 204
TYPE: prt
ORGANISM: Homo Sapien
IS-10-140-470-288

```

	Query Match	8.9%	Score 164.5;	DB 9;	Length 204;
	Best Local Similarity	27.7%	Pred. No. 4.1e-05;		
	Matches	46;	Conservative	31;	Mismatches 74; Indels 15; Gaps 5;
OY	18	LLGVTLVSGLSLEPVYWGNAANKRFQALEGGVLYYPQIGDRIDLICPRARPPGHS-S-PNY	76		
Db	8	LFGCCSIADRRATVTWSSNPKFRNE-DYTHVOLNDYDIIICFHYE---DISADAM	63		
OY	77	EFPYLIVYGAGRCRCEAPAPNILLICDRPDL--DLRFPIKFEQSYNNMGHFRSHH	133		
Db	64	EQYLLIVHEHEBYQLCPQSKXDYRWGQNPNPSAGHGPEKSERQRFTFFPLGLGEFKGH	123		
OY	134	DYLIANSDTGRGLESLGGVCITRGMKVLIRVGSPRGGAIVRK	179		

DB 124 SYYYISKPIHGHEDR-----CLRKLKTVSGKITHSPQAHNDPQE 162

RESULT 33

US-10-175-746-288

; Sequence 288, Application US/10175746

; Publication No. US200300227270A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: Deforge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P330R1C353

; CURRENT APPLICATION NUMBER: US/10/175,746

; PRIOR APPLICATION REMOVED - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 288

; LENGTH: 204

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-175-746-288

Query Match

Best Local Similarity 27.7%; Score 164.5; DB 9; Length 204;

Matches 46; Conservative 31; Mismatches 74; Indels 15; Gaps 5;

DB 18 LIGVLGVSGLSLEPYWNSANKRFGAGGYLYPQIGDRLLDLCPRARPPGHS--PNT 76

DB 8 LIGLCCSLAADRHVTFWNSSNPKFRNE-DYTHVQLNDYDILCPHYE--DHSADAM 63

QY 77 EFYKLYVGAQGRCEAPAPNLLTCDPDL--DLRFTIKFOEYSPNLMGHEFRSHH 133

DB 64 EGYILVIVHEHEVQLQPGSKQVWQCNRPKAKHGPEKLSKQRFPTFLGKEFEKGGH 123

QY 134 DYYIATSDGTRFEGLESLOGGYCLTRGKMKVLLRVGQSPRGCAVPRK 179

DB 124 SYYYISKPIHGHEDR-----CLRKLKTVSGKITHSPQAHNDPQE 162

US-10-176-918-288

; Sequence 288, Application US/10176918

; Publication No. US200300227275A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: Deforge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P330R1C288

; CURRENT APPLICATION NUMBER: US/10/176,921

; PRIOR APPLICATION REMOVED - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 288

; LENGTH: 204

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-176-921-288

Query Match

Best Local Similarity 27.7%; Score 164.5; DB 9; Length 204;

Matches 46; Conservative 31; Mismatches 74; Indels 15; Gaps 5;

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Zhang, Zemin

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P330R1C382

CURRENT APPLICATION NUMBER: US/10/176,918

Prior Application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 550

SEQ ID NO 288

LENGTH: 204

TYPE: PRT

ORGANISM: Homo Sapien

US-10-176-918-288

Query Match

Best Local Similarity 27.7%; Score 164.5; DB 9; Length 204;

Matches 46; Conservative 31; Mismatches 74; Indels 15; Gaps 5;

QY 18 LIGVLGVSGLSLEPYWNSANKRFGAGGYLYPQIGDRLLDLCPRARPPGHS--PNT 76

DB 8 LIGLCCSLAADRHVTFWNSSNPKFRNE-DYTHVQLNDYDILCPHYE--DHSADAM 63

QY 77 EFYKLYVGAQGRCEAPAPNLLTCDPDL--DLRFTIKFOEYSPNLMGHEFRSHH 133

DB 64 EGYILVIVHEHEVQLQPGSKQVWQCNRPKAKHGPEKLSKQRFPTFLGKEFEKGGH 123

QY 134 DYYIATSDGTRFEGLESLOGGYCLTRGKMKVLLRVGQSPRGCAVPRK 179

DB 124 SYYYISKPIHGHEDR-----CLRKLKTVSGKITHSPQAHNDPQE 162

US-10-176-921-288

; Sequence 288, Application US/10176921

; Publication No. US200300227276A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: Deforge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P330R1C288

; CURRENT APPLICATION NUMBER: US/10/176,921

; PRIOR APPLICATION REMOVED - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 288

; LENGTH: 204

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-176-921-288

Query Match

Best Local Similarity 27.7%; Score 164.5; DB 9; Length 204;

Matches 46; Conservative 31; Mismatches 74; Indels 15; Gaps 5;


```

; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and Pr
; FILE REFERENCE: DEX-0269
; CURRENT APPLICATION NUMBER: US/10/001,887
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/249,998
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/252,563
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 108
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-887-108

```

```

Query Match          7.5%; Score 138; DB 9; Length 638;
Best Local Similarity 27.9%; Pred. No. 0.011;
Matches 68; Conservative 13; Mismatches 79; Indels 84; Gaps 12;

```

```

QY 159 RCMKYLIVGQ-----SPRGAVPRKPYSEMPMERDGAASLERKENTLPGDPTSNATS 213
DB 137 RLKGDNGVQGPGLPAGAPQGGAPFP--GLP-----GPAGLGKPLDGLPAGAPQDGES 189
QY 214 -----RGAGPLPPSPMPAVA--GAAG-----G 234
DB 190 GPPGVPRGPRGAVGPPGPGVDGVPAAGALPGQSPGAGKGPETRGPLIGPTG 249
QY 235 LALLLL-----GYAGAGCAGCRRRRAKPSRHPG--PSSFGGSLGCGGGCKGPR 286
DB 250 YGMPGLPGKDRGAGVAGL---GDRGEDEDDPBGQPGGPGGPGGAGR 306
QY 287 EAPELGLALRG-----GAADPPCFHXKVSQDGYGHPIYVQGP--P 330
DB 307 RGPFPKGEAGCGPPGVPIRGDQPSGLAKKPGVPBERGLPGAHGP-----GPTGP 360
QY 331 QSPF 334
DB 361 KGEF 364

```

RESULT 40

```

US-09-823-240-10
; Sequence 10, Application US/09823240
; Patent No. US2002004813A1
; GENERAL INFORMATION:
; APPLICANT: James E. Bear
; APPLICANT: Frank B. Gertler
; APPLICANT: Jorgen Wehland
; APPLICANT: Joseph Loureiro
; TITLE OF INVENTION: Methods and Products for Regulating Cell
; FILE REFERENCE: M0656/7064 (HCL)
; CURRENT APPLICATION NUMBER: US/09/823,240
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/194,564
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-823-240-10

```

```

Query Match          7.1%; Score 130.5; DB 10; Length 380;
Best Local Similarity 22.8%; Pred. No. 0.022;
Matches 67; Conservative 34; Mismatches 102; Indels 91; Gaps 15;

```

```

QY 35 WMSAKRPAEGGYLYPIQIRLDLCPRRAPPPHSSPNEYKYLYVGAGRCCEA 94
DB 16 YDDGNRWLPAG--TGQAFSRVQI-----YHNPANSFRV-----GRKKQ- 55

```

```

QY 95 PPAENLLTCDBRDLDIRTIKQSESPY-----LMGHEFRSHHDYIIATSDGR 145
DB 56 -PQQQVINC-----ATVGVKYNQATPEFHQWRDARQVWGNFSGKEPAQFAA--GMA 107
QY 146 EGLESLOGVCVLTGKMLLRVQSPRG-----AVPRKPYSEMPMERD--GAASLER- 197
DB 108 SALEALEGG-----GPPPPALPTWASVNPSPSEVEBQQRQGPSEHIER 154
QY 198 -----PKENLPGDPTSN--ATSRGA--BGLPPSPMPAVA 229
DB 155 RVSNAGPPAPGAPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 214
QY 230 GAAGG-ALLLLGVAGAGCAGCRRRRAKPSRHPGPGSFGGSLGCGGGGM 283
DB 215 GREGGGA-----GAPGLAAIAGAKLR-KYSKQSEASGPTAKASGSGGGL 263

```

RESULT 41

```

US-09-925-299-979
; Sequence 979, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 979
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-979

```

```

Query Match          7.1%; Score 130.5; DB 10; Length 674;
Best Local Similarity 27.4%; Pred. No. 0.042;
Matches 72; Conservative 20; Mismatches 92; Indels 79; Gaps 15;

```

```

QY 141 SDGTREGLSLQGVCLTRGMKVLRLVQSP--RG--GAVPRKPYSEMPWE--RDRGAH 194
DB 186 SRGR-GLPGVAGV-----GEPPLGIAGPPGARGPAGVSGPVNAGAGGAGRDGNPGN 240
QY 195 SLERKENTRG-----DPTSNATSRGAEGLPP-----PMPAIVAGAA 232
DB 241 DGPFGDGGGKHGKRGYGNITGVGAGAPGPPGAVGAGHGRNGRTGPGP--VGR 298
QY 233 GGLALLLLGVAGAGCAGCRRRRAKPSRHPG--PSSFGGSLGL-----GGGG 281
DB 299 GA-----VGRPGSPGQGRGDKGEKPGGLGLKGNHGLQGLPIAGHHGPGAG 353
QY 282 GMP-----REARGELGA-LRG-----GGAADPPCFHXKVSQ 316
DB 354 SVGPAPRPRGAPGSPGAGDGTGHPGTVGPAGIRGPGQHGPGAPGPPGPPGPGVSG 413
QY 317 ---DYGHPIYIVQDPPQSPPN 336
DB 414 GGYDFGYDDDFYRADQPRASPSL 436

```

RESULT 42

```

US-09-854-133-187
; Sequence 187, Application US/09854133
; Publication No. US20020183495A1
; GENERAL INFORMATION:
; APPLICANT: Iodas, Michael J.
; APPLICANT: Mochamath, Raedoh
; APPLICANT: Henderson, Robert A.

```

```

; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 187
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-854-133-187

```

Query Match 7.0%; Score 129; DB 9; Length 595;

Best Local Similarity 27.4%; Pred. No. 0.0477; Mismatches 89; Indels 76; Gaps 12;

```

Db 163 GGAESPFGWEAGP-----REMGSPSGHGDGPRRRPRKRGK 201
Qy 128 EFRSHDYIITATSDGTREGLESLOGVCITRGMKVLRYGSPRGAVPRKPVSE---- 183
Db 202 RMGRHEAATAATATATGTGTAEAGASAPES-----QAGGPRGRA--RGPRQGRRR 254
Qy 184 MPMERDGAHSLPEPKENLPDPTS---NATSRGAGPLPPEPMFAVAGAGLALLL 240
Db 255 HGQRRRGPPQAREEG---PRDATTILGLTBSGEGRADQSGLPALAGAAAHAHAIP 310
Qy 241 G-----VAGAGAMCWRRRRAKPSRRHPGSGFGRGSLGLGGGGMGPREEAPGEL 293
Db 311 GAGPAAAPVGRGRGRGWRGGRGSGS---AGAGGGGRGR--GRGRGGGRG----- 357
Qy 294 GIALRGGAADP 305
Db 358 GGAAGGGGAAGP 369

```

```

Qy 294 GIALRGGAADP 305
Db 358 GGAAGGGGAAGP 369

```

```

RESULT 43
US-09-738-973-187
; Sequence 187, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodch
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Eliot, Mark
; APPLICANT: Mamion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 187
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-738-973-187

```

Query Match 7.0%; Score 129; DB 10; Length 595;

Best Local Similarity 27.4%; Pred. No. 0.0477; Mismatches 89; Indels 76; Gaps 12;

```

Qy 85 GGAAG-RCEAPAPNLLTCDRPDLRLFTIKQESPNLMGH----- 127
Db 163 GGAESPFGWEAGP-----REMGSPSGHGDGPRRRPRKRGK 201
Qy 128 EFRSHDYIITATSDGTREGLESLOGVCITRGMKVLRYGSPRGAVPRKPVSE---- 183
Db 202 RMGRHEAATAATATATGTGTAEAGASAPES-----QAGGPRGRA--RGPRQGRRR 254
Qy 184 MPMERDGAHSLPEPKENLPDPTS---NATSRGAGPLPPEPMFAVAGAGLALLL 240
Db 255 HGQRRRGPPQAREEG---PRDATTILGLTBSGEGRADQSGLPALAGAAAHAHAIP 310
Qy 241 G-----VAGAGAMCWRRRRAKPSRRHPGSGFGRGSLGLGGGGMGPREEAPGEL 293
Db 311 GAGPAAAPVGRGRGRGWRGGRGSGS---AGAGGGGRGR--GRGRGGGRG----- 357
Qy 294 GIALRGGAADP 305
Db 358 GGAAGGGGAAGP 369

```

```

RESULT 44
US-09-919-497-56
; Sequence 56, Application US/09919497
; Patent No. US2002010662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0601/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 56
; LENGTH: 1806
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (758)..(758)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: UNSURE
; LOCATION: (809)..(809)
; OTHER INFORMATION: Xaa = any amino acid
US-09-919-497-56

```

Query Match 7.0%; Score 129; DB 10; Length 1806;

Best Local Similarity 29.1%; Pred. No. 0.16; Indels 80; Gaps 13;

```

Qy 169 QSPRGAVPRPVSSEPMER-----DRGAHSLPEKKE-----NLRGD----- 206
Db 561 QGPRGVGPPGPGTGK-PGKGRPGADGGRGMPGPAKGRGFDGLGDKGHRGERG 619
Qy 207 -----PNSATSRGAGPLPPEPMFAVAGAGLALLLGVAGAGAMCWRRRRAKPS 260
Db 620 PQGPPGPGDDMGMDGEDEIGPRGLPGEABPRG-----LIGPRTPPA-----PQG 665
Qy 261 -----SRHPG-PGSFGRGSLGL-GGGGGMGPREEAPGELGIALRG 300
Db 666 PGMAGVDPGPKMNGMPQGPFGQGNPGPGGLGPGPGPIGP-----PGEKPGPKPG 721
Qy 301 -----GAADPPFCPHYEKVSGDYGHPIYIVODGP--PQSP 333
Db 722 LAGLPADGPPGHGKGGSGEKG-----ALGPPGPGP 755

```

Query Match 7.0%; Score 129; DB 10; Length 1806;

Best Local Similarity 29.1%; Pred. No. 0.16; Indels 80; Gaps 13;

Sequence 689, Application US/09925302

Patent No. US20020044941A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA104
CURRENT APPLICATION NUMBER: US/09/925,302
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05918
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 896
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 688
LENGTH: 403
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (183)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-689

Query Match 6.8%; Score 126.5; DB 10; Length 403;
Best Local Similarity 30.9%; Pred. No. 0.046;
Matches 46; Conservative 9; Mismatches 53; Indels 41; Gaps 8;

QY 198 PKEKULPDDPFSNMTS-----RGARGLPFPSPMPAVAGAGGLALLLLGVAGAGGCMWR 252
DB 168 PGAPGPPPPPEKAGXGAGRDGATGPGPGGPPGVKGEA-----LQPGGARG----- 217
QY 253 RRAKPSRRHGP-----GSGFRGSLGLGGGGMGPREAPGELGT-----ALRGGG 301
DB 218 ----KQAGTCTPGGCGGSGKSDGLTPKRTGT---TKGSKDLGLPGSKGRMKMGDA 270
QY 302 AADPPCPHYEKVSGDYGHPPYIVQDGP 330
DB 271 GVMGP--PGAQSGKDFGRP-----GPP 291

RESULT 46
US-09-978-295A-614
Sequence 614, Application US/09978295A
Patent No. US20020156006A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630P1C11
CURRENT APPLICATION NUMBER: US/09/978,295A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071

PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081195
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081203
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081229
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081955
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081817
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081819
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081952
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081838
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/082568
 PRIOR FILING DATE: 1998-04-21
 PRIOR APPLICATION NUMBER: 60/082569
 PRIOR FILING DATE: 1998-04-21
 PRIOR APPLICATION NUMBER: 60/082704
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082804
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082700
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082797
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082796
 PRIOR FILING DATE: 1998-04-23
 PRIOR APPLICATION NUMBER: 60/083336
 PRIOR FILING DATE: 1998-04-27
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/083392
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083495
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083496
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083499
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083445
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083554
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083558
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083559
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083500
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083742
 PRIOR FILING DATE: 1998-04-30
 PRIOR APPLICATION NUMBER: 60/084366
 PRIOR FILING DATE: 1998-05-05
 PRIOR APPLICATION NUMBER: 60/084414
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084441
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084637
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084639
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084640
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084598
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084627
 PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/084643
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/085339
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085338
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085323
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085582
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085700
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085689
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085580
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085573
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085704
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085697

Query Match 6.8%; Score 126.5; DB 9; Length 520;
 Best Local Similarity 30.3%; Pred. No. 0.061;
 Matches 46; Conservative 9; Mismatches 53; Indels 41; Gaps 8;

Qy 198 PGRNLTGDDPTSNMTS-----RGRBGLPPPSMRYAVGAGGJLALLLVAGAGGAGCMR 252
 Db 165 FGAPGPPPPPEAKAKAGMGDGTGSPGPPGVAGB-----LQGGQAGB----- 214
 Qy 253 RRRAPBSRRHGP-----GSFGGSLGCGGGMGFRAPBPLGT-----ALRGGG 301
 Db 215 ---KQATGTPGQGGKSGKDGGLGPKGEG---TKKEXDGLGPGSKGRMKMGDA 267
 Qy 302 AADPPFCHYKUSGDYGHVYIVQGGP 330
 Db 268 GWMGP--FGAQSGKDPGRP-----GPP 288

RESULT 47
 US-09-978-697-614
 Sequence 614, Application US/09978697
 Patent No. US20020169284A1
 GENERAL INFORMATION:
 APPLICANT: Ashkenazi, Avi
 APPLICANT: Baker Kevin P.
 APPLICANT: Botstein, David
 APPLICANT: Denoyez, Duc
 APPLICANT: Baton, Dan
 APPLICANT: Ferrara, Napoleon
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerber, Hanspeter
 APPLICANT: Geritsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, J. Christopher
 APPLICANT: Gurney, Austin L.
 APPLICANT: Hillan, Kenneth J.
 APPLICANT: Kijavlin, Ivar J.
 APPLICANT: Kuo, Sophia S.
 APPLICANT: Napier, Mary A.
 APPLICANT: Pan, James
 APPLICANT: Paoni, Nicholas F.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Shelton, David L.
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Williams, P. McKey
 APPLICANT: Wood, William I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
ACIDS
FILE REFERENCE: P2630PIC27
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078866
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08

PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627

PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 6.8%; Score 126.5; DB 9; Length 520;
Best Local Similarity 30.9%; Pred. No. 0.061;
Matches 46; Conservative 9; Mismatches 53; Indels 41; Gaps 8;

QY 198 PGKENTLPGDPTSNAT-----RGAEGPLPPSPMPAYAGAGLALLLVAGAGGAMCNR 252
DB 165 PGAPGPPEPAEKAKGKMGMRDGTGTPSGQGPVYKGEAG-----LQGPQAGP----- 214
QY 253 RRAKPSSESHPP-----GSFGRGSLGLGGGGMGPRBAEPGLGI-----ALRGGA 301
DB 215 ----KQANTGTBPQKSGSKSGDGLIGPKGTG---TKKEDLGLPSKGRGKMGKDA 267
QY 302 AADPFCHYKXGSDYGHFPTVQDGP 330
DB 268 GVMGP--PGAQSGKDGPR-----GPP 288

RESULT 48
US-09-978-192A-614
Sequence 614, Application US/09978192A
Patent No. US2002017753A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltzen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaudi, J. Christopher
APPLICANT: Guirney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavich, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630P1C9
CURRENT APPLICATION NUMBER: US/09/978,192A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049

PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 6.8%; Score 126.5; DB 9; Length 520;
Best Local Similarity 30.9%; Pred. No. 0.061;
Matches 46; Conservative 9; Mismatches 53; Indels 41; Gaps 8;

QY 198 RGENLRGDPSTNATSG--RGAEGPLRPMPVYVAGAGLALLLVNAGAGAMCR 252
DB 165 PGAPPPPPAEKXGKANGKRDATGSPGPPGVKXGAG-----LGGQAPG----- 214
QY 253 RRRAPSESRHPPG---GSFGRGSLGLGGGKNGPRRAEPGELGI-----ALRGCG 301
DB 215 ---KQATGTGPPQGEKSKDGGGILGPKGTG---TKGERKDILGSKKDRQWKQDA 267
QY 302 AADPFCHYKESGDYGHVYIVDDGP 330
DB 268 GVMGP--PGAQSKGDFGRP-----GFP 288

RESULT 49
US-09-999-832A-614
; Sequence 614, Application US/0999832A
; Publication No. US20020192706A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroli, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gertsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Guirney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel

APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PLC63
CURRENT APPLICATION NUMBER: US/09/999,832A
CURRENT FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08

PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600

;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 6.8%; Score 126.5; DB 9; Length 520;
Best Local Similarity 30.9%; Pred. No. 0.061;
Matches 46; Conservative 9; Mismatches 53; Indels 41; Gaps 8;

QY 198 PCKENLPGDPTSNATS-----RGARGLPFPSPMVAAGALMLLGVAGAGAMCMR 252
DB 165 PAAPGPGPAKAKAGKAMGRDAGTSGPGPGPVKKBAQ-----LGGPGAPG----- 214
QY 253 RRRAPPSERHPGPP-----GSFGRGSLGIGGGGGMGPRAEPGELGI-----ALRGGG 301
DB 215 ----KQAGTGTGPGGEEKSKGDDGLIGFKBTG---TKGEKGLDLGSKGDKGMKDA 267
QY 302 AADPPCPHYEVSGDYGHPIVVDGPP 330
DB 268 GVMGP--PGAQGSKGDFFGP-----GPP 288

RESULT 50
US-03-978-189-614
; Sequence 614, Application US/09978189
; Publication No. US20030004102A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gertsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.

;; APPLICANT: Tumas, Daniel
;; APPLICANT: Williams, P. Mickey
;; APPLICANT: Wood, William I.
;; TITLE OF INVENTION: Secured and Transmembrane Polypeptides and Nucleic
;; TITLE OF INVENTION: Acids Encoding the Same
;; FILE REFERENCE: P2630FIC7
;; CURRENT APPLICATION NUMBER: US/09/978,189
;; PRIOR FILING DATE: 2001-10-15
;; PRIOR APPLICATION NUMBER: 09/918585
;; PRIOR FILING DATE: 2001-07-30
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/064249
;; PRIOR FILING DATE: 1997-11-03
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066364
;; PRIOR FILING DATE: 1997-11-21
;; PRIOR APPLICATION NUMBER: 60/077450
;; PRIOR FILING DATE: 1998-03-10
;; PRIOR APPLICATION NUMBER: 60/077632
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077641
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077649
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; PRIOR APPLICATION NUMBER: 60/078004
;; PRIOR FILING DATE: 1998-03-13
;; PRIOR APPLICATION NUMBER: 60/078886
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078936
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078939
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079664
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079786
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079920
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/079923
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/080105
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080194
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070

PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 6.8%; Score 126.5; DB 9; Length 520;
Best Local Similarity 30.9%; Pred. No. 0.061;
Matches 46; Conservative 9; Mismatches 53; Indels 41; Gaps 8;

QY 198 PGKXNLTGDDPNSNATSGAEGPLPPSPMPAVVGAAGLALLLVGAAGGAMCWR 252
DB 165 PGADGPPPEPKAKAKAMGMDGATGSPGPGPVKXGNG-----LQGFQAGP----- 214
QY 253 RRAKPSRSRHPG-----GSFSGSLGAGGGGMPREAPGELGI-----ALRGGG 301
DB 215 ----KQATGTGPGQKSGKSGDGLGIPKGTG---TKEXKDLGLPGSKGDRGMGDA 267
QY 302 AADPPCPHIEKVSQDYGHPIYIVQDGP 330
DB 268 GVMGP--FGAGSKDPRG-----GPP 288

Search completed: February 11, 2003, 12:13:17
Job time : 20.5409 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 11:47:56 / Search time 25.2327 Seconds
(without alignments)
2776.398 Million cell updates/sec

Title: US-10-021-121-4

Perfect score: 1850
Sequence: 1 MGFPHSGPGVRRVALLILG.....PYIVQDGPQSPFNIVYKV 340

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database:

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phase:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriophage:*
17: SP_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1559	84.3	285	4 Q8TBH7	Q8TBH7 homo sapien
2	923	49.9	331	13 Q90Z31	Q90Z31 brachydanio
3	919	49.7	327	13 Q9PT69	Q9PT69 xenopus lae
4	629.5	34.0	333	13 Q9PU04	Q9PU04 gallus gall
5	616	33.7	341	13 Q90Z33	Q90Z33 brachydanio
6	605.5	32.7	334	13 Q90Z32	Q90Z32 brachydanio
7	340	18.4	205	13 Q9W6H9	Q9W6H9 xenopus lae
8	214.5	11.6	237	5 Q9U3M2	Q9U3M2 caenorhabdi
9	198.5	10.7	279	5 Q9V4E1	Q9V4E1 drosophila
10	182	9.8	652	5 Q9V4E1	Q9V4E1 drosophila
11	178.5	9.6	202	13 Q90Z31	Q90Z31 gallus gall
12	167	9.0	205	11 Q90Z38	Q90Z38 mus musculu
13	163.5	8.8	206	11 Q90Z38	Q90Z38 mus musculu
14	156	8.4	675	6 Q9N178	Q9N178 sus scrofa
15	155	8.4	229	13 Q9J431	Q9J431 brachydanio
16	153.5	8.3	1691	11 Q9ESQ2	Q9ESQ2 mus musculu

17	151.5	8.2	1447	13 Q9TB91	Q9TB91 xenopus lae
18	149	8.1	325	5 Q17036	Q17036 caenorhabdi
19	149	8.1	569	5 Q17208	Q17208 drosophila
20	148.5	8.0	316	5 Q19111	Q19111 caenorhabdi
21	146.5	7.9	1140	11 Q61434	Q61434 mus musculu
22	146.5	7.9	1491	13 Q91718	Q91718 xenopus lae
23	146.5	7.9	1774	11 Q62001	Q62001 mus musculu
24	146	7.9	675	13 Q90800	Q90800 gallus gall
25	146	7.9	1669	11 Q90Z50	Q90Z50 mus musculu
26	145.5	7.9	305	5 Q17805	Q17805 caenorhabdi
27	145.5	7.9	308	5 Q94620	Q94620 meloidogyne
28	145	7.8	680	11 Q9D0D2	Q9D0D2 mus musculu
29	145	7.8	1745	4 Q9N206	Q9N206 homo sapien
30	144.5	7.8	309	5 Q25466	Q25466 meloidogyne
31	144.5	7.8	1347	4 Q96093	Q96093 homo sapien
32	144.5	7.8	1420	13 Q90M37	Q90M37 gallus gall
33	143.5	7.8	775	16 Q9R342	Q9R342 streptomyces
34	143	7.7	1747	5 Q26640	Q26640 strongyloides
35	142	7.7	744	4 Q96D07	Q96D07 homo sapien
36	142	7.7	1835	13 Q91AV4	Q91AV4 gallus gall
37	141	7.6	284	5 Q25582	Q25582 teladorsagi
38	141	7.6	284	5 Q25581	Q25581 teladorsagi
39	141	7.6	319	5 Q17038	Q17038 caenorhabdi
40	141	7.6	809	13 Q93485	Q93485 oncorhynchus
41	141	7.6	1449	13 Q910C0	Q910C0 oncorhynchus
42	140.5	7.6	219	13 Q90YC5	Q90YC5 brachydanio
43	140.5	7.6	305	5 P91274	P91274 caenorhabdi
44	140	7.6	304	5 Q9XUE9	Q9XUE9 caenorhabdi
45	140	7.6	1378	5 Q97405	Q97405 halictus di
46	139.5	7.5	1450	13 Q91IB4	Q91IB4 cynops pyrr
47	139.5	7.5	1690	4 Q9UW66	Q9UW66 homo sapien
48	139.5	7.5	1691	4 Q9Y4L4	Q9Y4L4 homo sapien
49	139	7.5	705	4 Q8TEJ5	Q8TEJ5 homo sapien
50	139	7.5	744	11 Q9DZV4	Q9DZV4 mus musculu

ALIGNMENTS

RESULT 1	PRELIMINARY;	PRT;	285 AA.
Q8TBH7	Q8TBH7		
AC	Q8TBH7		
DT	01-JUN-2002 (TRENDS)		
DT	01-JUN-2002 (TRENDS)		
DT	01-JUN-2002 (TRENDS)		
DE	Similar to ephrin B3 (Fragment).		
OS	Homo sapiens (Human)		
OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Theria; Primates; Catarrhini; Hominoidea; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=BRAIN;		
RA	Strasbourg R.		
BL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC022499; AAH22499.1; -		
FT	NON_TER		
SQ	SEQUENCE	285 AA;	30159 MW; 08D8ACE7142469B2 CRC64;
Query Match	84.3%;	Score 1559;	DB 4; Length 285;
Best Local Similarity	100.0%;	Pred. No. 3.3e-112;	
Matches	285;	Conservative 0;	Mismatches 0; Gaps 0;
QY	56 DRDLICRRARPPGPHSGPNYFYKYLIVGAGQRCRCAPPAPNLLITCDRPPDILRFTI	115	
DB	1 DRDLICRRARPPGPHSGPNYFYKYLIVGAGQRCRCAPPAPNLLITCDRPPDILRFTI	60	
QY	116 KFOEISPLWGHFRSHDYITATSDTRGLESLLGGVCTREGMKVTLRVGQSPRGA	175	
DB	61 KFOEISPLWGHFRSHDYITATSDTRGLESLLGGVCTREGMKVTLRVGQSPRGA	120	
QY	176 VPRKVSMPMERDRGAHSLPEKENLPQDPTSNATSRGAEGLPPSPMPAVAGAGGL	235	

DB 121 VERKVSSEMPERDRGAHSLPEKXENLPQPTSNATRGAEPLPPSMVAVGAAGL 180
 QY 236 ALLLGVAGAGAMCWRRRRAKPSRRHPGPGSFGRGSLGLGGGGMGPPEAPGELGI 295
 DB 181 ALLLGVAGAGAMCWRRRRAKPSRRHPGPGSFGRGSLGLGGGGMGPPEAPGELGI 240
 QY 296 ALRGGGADPPFCPHYEKVSGDYGHPIYIVODGPPQSPNNIYKV 340
 DB 241 ALRGGGADPPFCPHYEKVSGDYGHPIYIVODGPPQSPNNIYKV 285

RESULT 2

Q90231 PRELIMINARY; PRT; 331 AA.
 ID 090231
 AC 090231
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Epirin B3.
 OS Brachdanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21290827; PubMed=11397014; Chen J.N., Goldstein N.B.,
 RA Chan J., Mably J.D., Serluca F.C., Chen J.N., Goldstein N.B.,
 RA Thomas M.C., Cleary J.A., Brennan C., Fishman M.C., Roberts T.M.;
 RT "Morphogenesis of prechordal plate and notochord requires intact
 RT eph/ephrin b signaling.";
 RL Dev. Biol. 234:470-482(2001).
 DR EMBL; AF375227; AAK64277.1; -.
 DR InterPro; IPR001799; Epirin.
 DR Pfam; PF00812; Epirin; 1.
 DR ProDom; PD002533; Epirin; 1.
 DR PROSITE; PS01299; EPHRIN; UNKNOWN_1.
 SQ SEQUENCE 331 AA; 35638 MW; 6A5BACD509A09818 CRC64;

Query Match 49.9%; Score 923; DB 13; Length 331;
 Best Local Similarity 54.8%; Pred. No. 3.1e-63;
 Matches 190; Conservative 43; Mismatches 74; Indels 40; Gaps 10;

QY 10 GVRVALLLLGLVLSLEPYVWNSANKRFQAEAGVYLYPQIGRLDLCRRAPRPG 69
 DB 9 GGLILFLVVDLG-ITATNMEPTVNSLNKRFSDKGYLYLPQIGRLDLCSSDPG 67
 QY 70 PHSPNVEFYKLYIVG-AQGRCEAPAPAPILLTCRDPDLRFTIKFQESPNLWGH 128
 DB 68 PAPAPADVEYKLYIVSSREQADRCVETGAPNLLTCDKPNSDMFTIKFQESPNLWGH 127
 QY 129 FSHHDYIATSDGTRBGLSLGGVCLTRGMKYLTVGSGPRG-GAVPRKPYSEMPME 187
 DB 128 FETNHDYIATSDGTRBGLSLGGVCLTRGMKYLTVGSGPRG-GAVPRKPYSEMPME 182
 QY 188 KDRGAHSLPEKXENLPQPTSNAT-----SRGAEGLPPSPMPAVAGAGLALL 239
 DB 183 -----AGRINNPNPGTGNSTHPOIPRSGSGENGPLPASINIAVIAAGAGSAPLL 232
 QY 240 LGVAGAGAMCWRRRRAKPSRRHPGPGSFG-----RGSLGCGGGGMGPPEAPGEL 293
 DB 233 L-VTVATVAVCTRRRAKPSRRHP-PLSLSLTSPKSGCGGAGGAGNNNG---SEPSDI 287
 QY 294 GIALGGGADPPFCPHYEKVSGDYGHPIYIVODGPPQSPNNIYKV 340
 DB 288 IPLR---TSDAYCPHYEKVSGDYGHPIYIVODGPPQSPNNIYKV 331

RESULT 3
 Q9PT69 PRELIMINARY; PRT; 327 AA.
 ID 09PT69
 AC 09PT69;

DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Epirin-B3 precursor.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Fipidae; Fipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TTSUTE-EMBRIO;
 RX MEDLINE=20099673; PubMed=10633856;
 RA Heibling P.M., Saulnier D.M.E., Robinson V., Christiansen J.H.,
 RA Wilkinson D.G., Brandt A.W.;
 RT "Comparative analysis of embryonic gene expression defines potential
 RT interaction sites for Xenopus EphA4 receptors with ephrin-B ligands.";
 RL Dev. Dyn. 216:361-373(1999).
 DR EMBL; AJ236866; CAB65511.1; -.
 DR InterPro; IPR002086; Aldehyde_dehydr.
 DR InterPro; IPR001799; Epirin.
 DR Pfam; PF00812; Epirin; 1.
 DR ProDom; PD002533; Epirin; 1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 DR PROSITE; PS01299; EPHRIN; 1.
 KW Signal.

FT SIGNAL. 1 24 POTENTIAL.
 SQ SEQUENCE 327 AA; 35913 MW; 4B80FA39D4C22DCD CRC64;

Query Match 49.7%; Score 919; DB 13; Length 327;
 Best Local Similarity 60.4%; Pred. No. 6.1e-63;
 Matches 192; Conservative 30; Mismatches 82; Indels 14; Gaps 8;

QY 25 VSGLSLEPYVWNSANKRFQAEAGVYLYPQIGRLDLCRRAPRPGPHSPNVEFYKLYIV 84
 DB 22 ISALSLDPTVWNSNKRFPDTEGYLYLPQIGRLDLCRRSPQGFSSPYEYKLYIV 81
 QY 85 GGAQG-RCEAPAPAPILLTCRDPDLRFTIKFQESPNLWGHFSHDYIATSDG 143
 DB 82 GTKEWSSCSILRTPLLLTCRDPDGLRFTIKFQESPNLWGHFGQSRDYIATSDG 141
 QY 144 TRBGLSLGGVCLTRGMKYLTVGSGPRGAVPRKPYSEMPERDRGAHSL-EPKXEN 202
 DB 142 TMDGIELTGGVCEYKGMKYLTVGSGPVGAPPRPSPS---AGDSGLSPSPNPDIPN 198
 QY 203 LPDPTSNATSGAEGLPPSPMPAVAGAGLALLLGVAGAGAMCWRRRRAKPSRR 262
 DB 199 V-GEISGNATKGENGFLLSHVPLVAGAGLALLL-VFGVGVCHRRQAKHSDTR 256
 QY 263 HPGPFGRGSLGLGGGGMGPPEAPGELGIALRGGGADPPFCPHYEKVSGDYGHPIV 322
 DB 257 HP-PLSLGSLTSPKSGNN-----NGHEPEDIIMPLRPSAAG--ACPHYEKVSGDYGHPIV 309
 QY 323 YIVODGPPQSPNNIYKV 340
 DB 310 YIVODMASQSPANNIYKV 327

RESULT 4
 Q9PU24 PRELIMINARY; PRT; 333 AA.
 ID 09PU24
 AC 09PU24
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Epirin-B2 precursor.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Menzel P., Pasquale E.B.;
 RT "coding sequence of chicken ephrin-B2";
 RL Submitted (Aug-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF180729; AAD53948.1; -
 DR InterPro; IPR001799; Ephrin.
 DR Pfam; PF00812; Ephrin.1.
 DR ProDom; PD002533; Ephrin.1.
 DR PROSITE; PS01289; EPHRIN; 1.
 KW Signal.
 FT SIGNAL.
 FT CHAIN
 SQ SEQUENCE 333 AA; 36761 MW; 4C28E8CB211B783 CRC64;

Query Match 34.0%; Score 629.5; DB 13; Length 333;
 Best Local Similarity 40.5%; Pred. No. 1.1e-40;
 Matches 135; Conservative 52; Mismatches 125; Indels 21; Gaps 6;

QY 14 GALLILGVLIVGSLSPYVNSANKRFQAEQGYLYPQIGRDLCLPRAPPSPHSS 73
 DB 16 GALLVLRRLAKSIVLDPYVNSNPKFLPGGLVLYPQIGKDLICPKV---DSKTV 72
 QY 74 PNEYFYLYVGAQGRCAAPAPNLLTCDPDLRTTFQGYSPNLMGHEFRSH 133
 DB 73 GQTEYKVVWVDKQADSCAIRKNTPELLNCARPDDVKTTFKFQFSPLNGLEFQKNK 132
 QY 134 DYIIATSDGTREGLESLOGVCLTRGMKVLRVGSPRGAVPRKPVSEMPER----- 188
 DB 133 DYVISTNSGSLGANNQEGVGCOTKTKLMVGGDPPNSAGLPR---STDPIKREQQA 189
 QY 189 -DRGAHSLPEQENLPDPTSNATRGAEGLPPSPMPAVAGAGLALLLVAGAGG 247
 DB 190 GTNGKSTTSPFYKDHSGSSTDS--SKAGHSILGSEVALFAGIASGCTIFVIITLVV 247
 QY 248 AMCRRRRAPKSESRHPGSGFRGSGSLGCGGGMGPRAEPGEIGLIRGGADPPF 307
 DB 248 LLLKTRRRKSPQHTTILSLTATPKRSGNN---NGSEPSDIIFLR---TADSVF 300
 QY 308 CPYKESGDIYGVYIVDGPQSPSPNITYKV 340
 DB 301 CPYKESGDIYGVYIVDGPQSPSPNITYKV 333

RESULT 5
 ID Q90233 PRELIMINARY; PRT; 341 AA.
 AC Q90233;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Ephrin B1.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NC NCB1_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21290827; PubMed=11397014;
 RA Chan J., Mably J.D., Serluca F.C., Chen J.N., Goldstein N.B.,
 RA Thomas M.C., Cleary J.A., Brennan C., Fishman M.C., Roberts T.M.;
 RT "Morphogenesis of prechordal plate and notochord requires intact
 RT eph/ephrin b signaling."
 RL Dev. Biol. 234:470-482(2001).
 RL EMBL; AF375224; AAK64274.1; -
 DR InterPro; IPR001799; Ephrin.
 DR Pfam; PF00812; Ephrin.1.
 DR ProDom; PD002533; Ephrin.1.
 DR PROSITE; PS01289; EPHRIN; 1.
 SQ SEQUENCE 341 AA; 37849 MW; CB9222F20B0D93E94 CRC64;

Query Match 33.3%; Score 616; DB 13; Length 341;
 Best Local Similarity 41.8%; Pred. No. 1.3e-39;
 Matches 141; Conservative 49; Mismatches 109; Indels 38; Gaps 11;

QY 24 LVGSLSPYVNSANKRFQAEQGYLYPQIGRDLCLPRAPPSPHSSPNYEFYKYL 83
 DB 23 LPAKSLSEVWNSSQNPFXVSGKGLIVYPEIGDKLDITCPK-----GDMGRD-YEFYKYL 77
 QY 84 VQGAQGRCAAPAPNLLTCDPDLRTTFQGYSPNLMGHEFRSHDYIIATSDG 143
 DB 78 VKGQAECSCTILDPNVLYTNCKPEKDKTKTFKFQFSPLNGLEFQKFTVYITTSNG 137
 QY 144 TRGLESLOGVCLTRGMKVLRVGSPRG-----GAVPRKPVSEMPERDRGAHSLSP 198
 DB 138 TOGLENNEGVCSTRSMKIMKVGQDPNADPDLPLDPRPYNEIKDPTSPSRKTER 197
 QY 199 GKEN-----LPQDPSNATSR--GAEGLPPSPMPAVAGAG--GIALLLGVAGAG 247
 DB 198 GRENEVDNGSRMKGKTRNQNNSPGVEGIF--GSXPALFAAGACVFLIILIV 255
 QY 248 AMCRRRRAPKSESRHPGSGFRGSGSL---GLGGGGMGPRAEPGEIGLIRGGGA 303
 DB 256 LLLKTRKTR---KHSQ---RGTLASLTATPKRGAAGASPSDITIFLR---TT 304
 QY 304 DPPCPHYEKVSDYGVYIVDGPQSPSPNITYKV 340
 DB 305 ENNYCPHYEKVSDYGVYIVDGPQSPSPNITYKV 341

RESULT 6
 ID Q90232 PRELIMINARY; PRT; 334 AA.
 AC Q90232;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Ephrin B2b.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NC NCB1_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21290827; PubMed=11397014;
 RA Chan J., Mably J.D., Serluca F.C., Chen J.N., Goldstein N.B.,
 RA Thomas M.C., Cleary J.A., Brennan C., Fishman M.C., Roberts T.M.;
 RT "Morphogenesis of prechordal plate and notochord requires intact
 RT eph/ephrin b signaling."
 RL Dev. Biol. 234:470-482(2001).
 RL EMBL; AF375226; AAK64276.1; -
 DR InterPro; IPR001799; Ephrin.
 DR Pfam; PF00812; Ephrin.1.
 DR ProDom; PD002533; Ephrin.1.
 SQ SEQUENCE 334 AA; 36998 MW; 341497E0F82472BD CRC64;

Query Match 32.7%; Score 605.5; DB 13; Length 334;
 Best Local Similarity 41.0%; Pred. No. 7.9e-39;
 Matches 134; Conservative 45; Mismatches 113; Indels 35; Gaps 8;

QY 30 LEFYVNSANKRFQAEQGYLYPQIGRDLCLPRAPPSPHSSPNYEFYKYLVGAGG 89
 DB 27 LESVYNTSNKFPVGGVLYPQIGKDLICPKV---PSTGQTNLEFRVLYVPEQL 85
 QY 90 RRCAPAPNLLTCDPDLRTTFQGYSPNLMGHEFRSHDYIIATSDGTREGLE 149
 DB 86 ETCVTVSDMLLNCDCRDPDVKTFKFQFSPLNGLEFRKQDVYIISTNSSTPGLD 145
 QY 150 SLOGVCLTRGMKVLRVGSPRGAV---PRKPVSEMPERDRGAHSLPEGKEVLP 205
 DB 146 NHHGQVCRSKSMKVLRVGSPPTDSBSAKNHPTRNPKYIENKQDNTF-----SKENDVS 200
 QY 206 --DPTSNATRGAEGLPPSPMPAVAGAGLALLLVAGAGAGCW-----RR 253
 DB 201 QIDSMQNGESGSGXSG-----ESVSGASDVALLFAGV--AGAVIFLIILVALLH 250

QY 254 REAKESRHPGSGFRGSGSLGCGGCMGPBAPGELGALRGGAADPFCPHYK 313
 DB 251 RRHQGHSAQCSCQLPLNTLPKRGSGASGSGNNNGSEPDIFPIRTSGM---YCPHYK 307
 QY 314 VSGDYGHPPYIVQDGPSPNNIYKV 340
 DB 308 VSGDYGHPPYIVQEMPONPANITYKV 334

RESULT 7

Q9W6H9 PRELIMINARY; PRT; 205 AA.
 AC Q9W6H9;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Ephrin-B2 (Fragment).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus.
 NC NCB1_TaxID=83355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97411149; PubMed=9259557;
 RA Smith A., Robinson V., Patel K., Wilkinson D.G.,
 RT "The EphA4 and EphA1 receptor tyrosine kinases and ephrin-B2 ligand
 RT regulate targeted migration of branchial neural crest cells.",
 RL Curr. Biol. 7:561-570(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20093673;
 RA Hebling F.M., Saulnier D.M., Robinson V., Christiansen J.H.,
 RA Wilkinson D.G., Brandt A.W.,
 RT "Comparative analysis of embryonic gene expression defines potential
 RT interaction sites for Xenopus EphA4 receptors with ephrin-B ligands",
 RL Dev. Dyn. 216:361-373(1999).
 DR EMBL; AF128844; AAD32610.1;
 DR InterPro; IPR001799; Ephrin.
 DR Pfam; PF00812; Ephrin; 1.
 DR ProDom; PD002533; Ephrin; 1.
 FT NON TER 1
 SQ SEQUENCE 205 AA; 22256 MW; 7DEDD34277260C87 CRC64;

Query Match 18.4%; Score 340; DB 13; Length 205;
 Best Local Similarity 37.0%; Pred. No. 1.2e-18;

Matches 84; Conservative 35; Mismatches 72; Indels 36; Gaps 7;

QY 128 EFRSHDYIATSDGTREGLSLQGVCLTRGMKVLRLVQSP-----RCGAVPRKPV 182
 DB 1 EFGRRKDYIISTNGSLGVNCGVCYTKMKILMKVQDPNFMHNRGASSTRPDH 60
 QY 133 EYPM--EEDRGAAHSLFEGKENTPEDPTSNATSGAEGPLPPSPMPAVGAAGLALLL 240
 DB 61 ESGTNGKSTTSPHVGEGSSTEGKNAGHSIIIGSEVAL-----FAGIASIIIFIV 113
 QY 241 GVAAGAGMCMRRRAKPSSESRHPGSGFRGSGSLG-----CGGGMPPRAEPGEL 293
 DB 114 IITIVLILKTRRRHRKHSPOH-----TISLSTLTPKSGNNNG---SEPSDI 161
 QY 234 GIALRGGAADPFCPHYKXVSGDYGHPPYIVQDGPSPNNIYKV 340
 DB 152 IIPLR---TAGGVFCPHYKXVSGDYGHPPYIVQEMPONPANITYKV 205

RESULT 8

Q9U3M2 PRELIMINARY; PRT; 237 AA.
 AC Q9U3M2;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE C43F9.8 protein.

GN C43F9.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Felodierinae; Caenorhabditis.
 NC NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mortimore B.J.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology.",
 RL Science 282:2012-2018(1998).
 DR EMBL; Z62252; CAB54195.1;
 DR InterPro; IPR001799; Ephrin.
 DR Pfam; PF00812; Ephrin; 1.
 DR ProDom; PD002533; Ephrin; 1.
 SQ SEQUENCE 237 AA; 26748 MW; B9B2D9FCCT1FE4FC CRC64;

Query Match 11.6%; Score 214.5; DB 5; Length 237;
 Best Local Similarity 25.7%; Pred. No. 6.1e-09;

Matches 53; Conservative 41; Mismatches 83; Indels 29; Gaps 6;

QY 11 VRVGAALLGLVGLVS-GSLPEVYVNSANKRFOAG-GVLYYPOIGRLDLCPRAPP 68
 DB 1 MQATITLTLSPFLGMARKIPIDIMWISNPIFDVSNMDHVSVHIGRVSIRCKSDPT 60
 QY 69 GPHSSPNYEFKLYLVGAGRCRCEAPPAFNNLLTCDDPDDLTFTIYQESPLNKGH 128
 DB 61 G-----KYEYSYTYVWSDEYDHCFL-SKPRLVGACDQTNASINIVFRSSTPPGGE 114
 QY 129 FRSRHDYIIA-----TSDGTREGLSLQGVCLTRGMKVLRLVQ 169
 DB 115 FQGGKMYFLISSEVDALIVETANQIFPGTSDGLGIDRGDGLCAKCKIKIFVQ 174
 QY 170 SPFGAVPRKPVSEMERDGAHS 195
 DB 175 DRGIENPK--FAARTLKDRDAHS 198

RESULT 9

Q9U474 PRELIMINARY; PRT; 279 AA.
 AC Q9U474;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE VAB-2 (Hypothetical protein Y37E11AR.6).
 GN VAB-2 OR Y37E11AR.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 NC NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N2;
 RX MEDLINE=20084449; PubMed=10619431;
 RA Chin-Sang I.D., George S.E., Ding M., Moseley S.L., Lynch A.S.,
 RA Chisholm A.D.;
 RT "The ephrin VAB-2/EFN-1 functions in neuronal signaling to regulate
 RT epidermal morphogenesis in C. elegans.",
 RL Cell 99:761-780(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.",
 RL Science 282:2012-2018(1998).


```

DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE Ephrin-A6 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauira; Aves; Neognathae; Galliformes; Psittacidae; Psittacinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Menzel P., Valencia F., Godement P., Dodelt V.C., Pasquale E.B.,
RT "Ephrin-A6, a new ligand for EphA receptors in the developing visual
system."
RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF317286; AAK0944.1;
DR InterPro; IPR01799; Ephrin.
DR Pfam; PF00812; Ephrin.1.
DR PRINTS; PR01347; Ephrin.1.
DR PRODOM; PD002533; Ephrin.1.
DR PROSITE; PS01299; EPHRIN; 1.
DR NON TER
SQ SEQUENCE 202 AA; 22624 MW; 91E2716FP9135F9 CRC64;

Query Match 9.6%; Score 178.5; DB 13; Length 202;
Best Local Similarity 27.5%; Pred. No. 3e-06;
Matches 56; Conservative 27; Mismatches 84; Indels 37; Gaps 7;

QY 33 YVWNSANKRFGQEGVYVYPOIGDRDLCPARPSPGHSPPNYFYLVGAGQGRGRC 92
DB 25 YVWNSNRRF-LQDDYSLOVSIINDHDIYCHYSAPTMA-----ESFLFWDEGRGRC 79
QY 93 EAPPAPLLITCDR--PDLDRTFKFOEYSPNLWGHFESHHDYTIAT-SDGTSEGL 148
DB 80 SETPAFAKMECNKPPAPFVPVPRFEKIQRPFSGLGFEFFGERTYYYSVPTFGS---- 135
QY 149 ESLQGVCLTRGKMTLVKVGSPRGAVPRKVPSEMPERDRGAHSLPEKENLPQDPT 208
DB 136 ---AGRLCKRVSYCCR-----ASTPEPLTEVPNSQRRG-----GQPE 171

QY 209 SNATSRGAGEPPLPPSPMAVAGAA 232
DB 172 GDGSPRDAAFPQRSRTLVALA 195

RESULT 12
Q9D7K8 PRELIMINARY; PRT; 205 AA.
AC Q9D7K8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:331004015, full insert sequence.
GN EFNA1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojopori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Iono H., Baldarelli R., Barsh G.,
RA Blake U., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,

```

```

RA Brownstein M.J., Bulb C., Fletcher C., Fujita M., Gariboldi M.,
RA Guistinich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima U., Mazarelli U., Momberts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlsuki S.,
RA Hasehizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK009144; BAB26102.1;
DR MGD; MGI:103236; Etnal.
DR InterPro; IPR001799; Ephrin.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00812; Ephrin.1.
DR PRINTS; PR01347; EPHRIN.
DR PRODOM; PD002533; Ephrin.1.
DR PROSITE; PS01299; EPHRIN; 1.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN 1.
DR PROSITE; PS00678; WD_REPEATS_1; E37E55767459A4BC CRC64;
SQ SEQUENCE 205 AA; 23772 MW; E37E55767459A4BC CRC64;

Query Match 9.0%; Score 167; DB 11; Length 205;
Best Local Similarity 26.1%; Pred. No. 2.3e-05;
Matches 43; Conservative 34; Mismatches 76; Indels 12; Gaps 4;

QY 18 LIGVLVGLVGLSLPEYVWNSANKRFGQEGVYVYPOIGDRDLCPARPSPGHSPPNYE 77
DB 8 LIGLCCSLAADRHIVFNWNSNPKFREE-DYTVHQLNDYLDICHYVEDSDV-ADAAVE 65
QY 78 FYLVYVGAQGRRCRCPAPPAPPLLITCDRDL--DLRTFKFOEYSPNLWGHFESHHD 134
DB 66 RYLVVWHEHVEYVACOPQSKDOVRNCRNPSAKHGPEKLSVFORPTFLIKERFEGHS 125
QY 135 YXLIATSDGTRREGLESLQGVCLTRGKMTLVKVGSPRGAVPRK 179
DB 126 YXXISKPIYHGE-----SQCCLKVYVNGKRTNPNQVHNPQE 163

RESULT 13
Q9CZS8 PRELIMINARY; PRT; 206 AA.
AC Q9CZS8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 10 days embryonic cDNA, RIKEN full-length enriched library,
DE clone:2610529M21, full insert sequence.
GN EFNA4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojopori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Iono H., Baldarelli R., Barsh G.,
RA Blake U., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bulb C., Fletcher C., Fujita M., Gariboldi M.,
RA Guistinich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima U., Mazarelli U., Momberts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlsuki S.,

```

RA Hayashizaki Y.;
 RT "functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK012195; BAB28092.1; -
 DR MGI; MGI:106643; Efn4.
 DR InterPro; IPR001799; Ephrin.
 DR Pfam; PF00812; Ephrin.1.
 DR PRINTS; PR01347; EPHRIN.
 DR ProDom; PD002533; Ephrin.1.
 DR PROSITE; PS01299; EPHRIN.1.
 SQ SEQUENCE 206 AA; 22859 MW; 675E32971D1C6EBC CRC64;

Query Match 8.8%; Score 163.5; DB 11; Length 206;
 Best Local Similarity 28.1%; Pred.No.4.3e-05;
 Matches 61; Conservative 17; Mismatches 80; Indels 59; Gaps 10;

QY 32 PYVNSANKRFOAGGYLYPQIGRLDLICPRARPPGPHSSPNVEFYKLYLVGAQGR 91
 DB 29 PIYVNSNPRL-LRGDAVVELGENDYLDIFCPHYESPGE--ETFALYIVMSGYEA 85
 QY 92 CEAPPANL-LITCDRPLD---DLRTIKQGYSPNLMGHERSHHDYITATSDGTREG 147
 DB 86 CTAEGANSPQRMNSMPAPFSPVSEKIQRTPLPLGFEFLPGETYYISVPTPESPG 145
 QY 148 -LESLOGVCLTRGMKYLRLVQSPRGAVPRKPVSEMERDRGAH-SLEPKENLP 205
 DB 146 RCLRLGVVCC-----KESGSHSHP-----VG 170
 QY 206 DPTSNATS--RGAEGLPPSPMPAVAGAGLALLL 240
 DB 171 SPESGTSQWKGHPSP-----LCLLL 194

RESULT 14

Q9N178 PRELIMINARY; PRT; 675 AA.

AC Q9N178; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Type X collagen.
 CN COL10A1.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Suidae.
 NCBI_Taxid=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21015405; PubMed=11130976;
 RA Nielsen V.H., Bendixen C., Arndbjerg J., Sorensen C.M., Jensen H.E.,
 RA Shukri N.M., Thomsen B.;
 RT "Abnormal growth plate function in pigs carrying a dominant mutation
 in type X collagen."
 RT Mamm. Genome 11:1087-1092(2000).
 RL EMBL; AF22861; AAF37271.1; -
 DR InterPro; IPR001073; C1Q.
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF00386; C1Q.1.
 DR Pfam; PF01391; Collagen; 8.
 DR PRINTS; PR00007; COMPLEMENTC1Q.
 DR ProDom; PD000007; collagen; 2.
 DR SMART; SM00110; C1Q.1.
 DR PROSITE; PS01113; C1Q.1.
 DR Collagen.
 SQ SEQUENCE 675 AA; 65447 MW; 26397B10310363F9 CRC64;

Query Match 8.4%; Score 156; DB 6; Length 675;
 Best Local Similarity 28.6%; Pred.No.0.00063;
 Matches 106; Conservative 17; Mismatches 118; Indels 130; Gaps 26;
 QY 1 MGPFHSGPGCVRYGALLLGLVGLSLPYYNSANKRFOAGGYLYPQI-GDR-L 58
 DB 212 MGPP--GPPGV-----GKR--GENGPGQGIKDRGF 240

QY 59 DLICPRARPPGPHSSPNVEFYKLYLVGAQGR-CEAPPANLITCDRPLDLRTIKF 117
 DB 241 PGESGPAQPPGPGP-----GEGRBEIGRGAAG--AAGQGL----- 278
 QY 118 QEYSPNLMGHERSHHDYITATSDGT---RGLSELQGVCLTRGMKYLRLVQSPRG 173
 DB 279 ---PGTKGHPGAG---YAGPPGAPGPGKPLPLKQ---QRG-----PIG 315
 QY 174 GAVPRKPVSEMERDRG-AAHSLPEKKNLPDPTSNATSRGAEGLPPSPVAVGAA 232
 DB 316 --LPGAGG---KGEQGPAGHPPEPLTGP--SRGQSPKGLPANNVGP 362
 QY 233 GGLALLLVAGAGAGACWRRR---AKPSESHPG-----PGSTGRGSLGLGGGG 282
 DB 363 GEIG--LAGPAGPPGAGXGERGPGSLDGPVGPFGPIGPKGNPLGPXKDDPG 420
 QY 283 M---GPREA--PGEIGL-LRG-----GAADPPCPHYEKVSGDYPIYIVQD 328
 DB 421 LPGVGPAGAKVGNHNGEAGPRGAPGIPGTRGPIGPPIGFPGSKDGNP-----G 474
 QY 329 PPGSPNIYK 339
 DB 475 PP-GPAGIATK 484

RESULT 15

Q93431 PRELIMINARY; PRT; 229 AA.

AC Q93431; 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Ephrin A-1.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NCBI_Taxid=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Durbin L., Brenan C.H., Shiomi K., Cooke J.;
 RT "Eph signalling is required for segmentation and differentiation of
 the somites."
 RT Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RL EMBL; A006838; CA0764.1; -
 DR InterPro; IPR001799; Ephrin.
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF00812; EPHRIN.1.
 DR PRINTS; PR01347; EPHRIN.
 DR ProDom; PD002533; Ephrin.1.
 DR PROSITE; PS01299; EPHRIN.1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN 1.
 SQ SEQUENCE 229 AA; 26115 MW; 8664462F67AF6F5C CRC64;

Query Match 8.4%; Score 155; DB 13; Length 229;
 Best Local Similarity 27.2%; Pred.No.0.00022;
 Matches 56; Conservative 23; Mismatches 93; Indels 34; Gaps 7;

QY 33 PYVNSANKRFOAGGYLYPQIGRLDLICPRARPPGPHSSPNVEFYKLYLVGAQGR 92
 DB 25 VYVNSANFLMD-DYVDVRLNDYDITCPH-YANGELISGAELRYLVWLEEDYENC 82
 QY 93 EAPPANLITCDR---FDLRLFTIKQGYSPNLMGHERSHHDYITATSDGTREG 149
 DB 83 KPHSFQDLRWECGRPFAPFAPEKSEKQRTPLPLGFEFLPGETYYISVPTPESPG 145
 QY 150 SL--QGVCITRGMKYLRLVQSPRGAVPRKPVSEMERDRGAH-SLEPKENLP 207
 DB 134 PLHHGQECRLKVDV---GPHGSKKKQKQVVEETEGTAAAGVYNNPNNRLPAD 188
 QY 208 TSNATSRGAEGLPPSPMPAVAGAG 233

Db 189 -----PIAMIPVQBSVG 201

RESULT 16

Q9ESQ2 ID Q9ESQ2 PRELIMINARY; PRT; 1691 AA.
 AC Q9ESQ2;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Type IV collagen alpha 5 chain.
 GN COL4A5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
 NC NCB1_TaxID=10090;
 RN [1]
 RX MEDLINE=20536494; PubMed=10965041;
 RA Saito K., Naito I., Seki T., Ohashi T., Kimura E., Momota R.,
 RA Kishimoto Y., Sado Y., Yoshitaka H., Nishimura Y.,
 RT "Differential Expression of Mouse a5(IV) and a6(IV) Collagen Genes in
 Epithelial Basement Membranes";
 RL J. Biochem. 128:427-434(2000).
 DR EMBL; AB041350; BAB13673.1; -.
 DR MGD; MGI:88456; Col4a5.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR001442; Procollagnc4.
 DR Pfam; PF01413; C4; 2.
 DR Pfam; PF01391; Collagen; 24.
 DR ProDom; PD000007; Collagen; 3.
 DR ProDom; PD003923; Procollagnc4; 1.
 DR SMART; SM00111; C4; 2.
 KW Collagen.
 SQ SEQUENCE 1691 AA; 161823 MW; 81340DF1792208FA CRC64;
 Query Match 8.3%; Score 153.5; DB 11; Length 1691;
 Best Local Similarity 24.6%; Pred. No. 0.0028;
 Matches 86; Conservative 29; Mismatches 119; Indels 115; Gaps 18;
 QY 52 PGIQRLDLPPRRAPPGPHSSPYVEFYKLYLVGAGRCCEAPAPNLLTCDRPPDL 111
 Db 442 PQLPSR-DEIC-KAGPPPPPPGDK-----GLQEGGVKGDCKDTPKNC----- 484
 QY 112 RFTIKFOEYSPNLWGHFRSHHDYIATSDTRREGLESLOG-----GVCLTRGMKV 163
 Db 485 -----IGTGISPPGQGLPGLGPPGSLGIPGKGDK- 517
 QY 164 LLRVGQS---PRG-GAVPRKP-----VSEMP-MERDRG-----AAHSL- 196
 Db 518 ---GQAQITGPKLPGLPFPFGAPGFGSGKGDGDLTLPGMGDGLGELGFPGLPG 573
 QY 197 ---EPGKENTLPDPP---TSNATSGAGGLPFPSPMPAVAGAGLALL----- 238
 Db 574 LPGSGKXGGLGLPGPKPGGIGITFKGERPGRGSLPGLPGNNKGPFGPFGPPTGE 633
 QY 239 --LLGVAGAGACMKRRRAKPSRS-RHFG---PGSGRGSGSLGLGGGGMGRRAEPG 291
 Db 634 KGIQGVANPQPGLPGRKPGGQITTOPGRGLPGSGRGEVGLGDLPLPQGLPG 693
 QY 292 ELGI---ALRGGAADPPCFPHYEKVSQD---YGHPIYIVQDGPPOSP 333
 Db 694 ILGSGEGGIGIPIGPBPGRPKGFGIIRPGARAGARMGDPDPGPP 742
 RESULT 17
 Q9IB91 ID Q9IB91 PRELIMINARY; PRT; 1447 AA.
 AC Q9IB91;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Type I collagen alpha 1.

COL1A1.

GN COL1A1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus.
 NC NCB1_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Goto T., Katada T., Kinoshita T., Kubota H.Y.;
 RT "Expression and characterization of Xenopus type I collagen alpha 1
 (COL1A1) during embryonic development";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB034701; BAA94972.1; -.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR000885; Fib collagen_C.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF01410; COLFI; 1.
 DR Pfam; PF01391; Collagen; 18.
 DR Pfam; PF00093; VWC; 1.
 DR ProDom; PD000007; Collagen; 2.
 DR ProDom; PD002078; Fib collagen_C; 1.
 DR SMART; SM00038; COLFI; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01208; VWF_C; 1.
 KW Collagen.
 SQ SEQUENCE 1447 AA; 137446 MW; AAA6DD2B4158E38B CRC64;
 Query Match 8.2%; Score 151.5; DB 13; Length 1447;
 Best Local Similarity 27.2%; Pred. No. 0.0034;
 Matches 86; Conservative 13; Mismatches 128; Indels 89; Gaps 16;
 QY 67 PPGHSSPYVEFYKLYLVG-----GAGRCCE-APPAPNLLTCDRPPDL 111
 Db 756 PPGPGAGPDK-----GEAGPAGPAGTARAGAPGERGEPAPPGAGF--AGPPAD- 805
 QY 112 RFTIKFOEYSPNLWGHFRSHHDYIATSDTRREGLESLOG-----GVCLTRGMKV 162
 Db 806 ---GQPAKKEQSDG---AKDAGPGLGPGTGPAPGALSPGKARG 851
 QY 163 -----VLLRVG-QSPFGAVPRKPVSEMPMERDRG-----AAHSLPGEKEN 202
 Db 852 AGPPGATGFPAGARVPPPSGNAGPPGSPGAKGKAGAPGRTGPAGSGEPGAAG 911
 QY 203 LPGDPTSNATSGAGGLPFPSPMPAVAGAGLALL-LLGVAGAGACMKRRRAKPSRS 261
 Db 912 PPGPGEKG-SFGSDGPPAGAPGIPGPGVAGSKGTGVLPGKRGERSGGLPGPAGEPKQ 970
 QY 262 RHPRP---GSPGRGSGSLGLGGGGMGRRAEPGELGIALRGGAADPPCFPHYEKVSQD 317
 Db 971 GSGSGSERGPPGPPGLGPPGSGRGAAPSGEAPGR-DGAVP-----KGD 1020
 QY 318 YGHPIYIVQDGPPOSP 333
 Db 1021 RGEAA-----GPPGAP 1031
 RESULT 18
 Q17036 ID Q17036 PRELIMINARY; PRT; 325 AA.
 AC Q17036;
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Hypothetical 31.6 kDa protein.
 GN T15B7.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 NC NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 MEDLINE=99069613; PubMed=9851916;

RA None;
 RT "Genome sequence of the nematode *C. elegans*: a platform for
 RT investigating biology. The *C. elegans* Sequencing Consortium.";
 RL Science 282:2012-2018 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Pauley A., Gattung S.;
 RT "The sequence of *C. elegans* coemid T15B7.";
 RT Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF022885; AAB69961.1; -
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR002486; Col cuticle_N.
 DR Pfam; PF01391; Collagen_3.
 DR Pfam; PF01484; Col cuticle_N; 1.
 DR Hypothetical protein.
 KW SEQUENCE 325 AA; 31591 MW; 6538D4537C9EC202 CRC64;
 SQ
 Query Match 8.1%; Score 149; DB 5; Length 325;
 Best Local Similarity 26.0%; Pred. No. 0.00095;
 Matches 92; Conservative 25; Mismatches 109; Indels 128; Gaps 23;

QY 13 VGALLILGVLGSL-----EPVYNSANKRFQAEQ-----VLYPQIGDRLLDLIC 62
 DB 1 MSASTVTVASAAGIAIVCFVTGMIFFINDINSFDEKIGEKKEKYEQLA-WQAMI 58
 QY 63 PRARPPPHSSPYEFYKLVNG-----GAGRRCFA-PPRPNLLTCDRDL 111
 DB 59 PTPRPSSGSS-----PLLRNKRQACNCAGSRGCPAPRPP 96
 QY 112 RTIKFQEVSPNLMGHEFRSHNDYIIATSDTREGLESLOG--GYCLTRGMKVLRLVQ 169
 DB 97 -----PGQPGAPGEGHNG-----LAQPGSGARINPATGPGFCIT----- 133
 QY 170 SPFGAVPRKPVSEMPERDRGAHSLERKENTLPGDPTSNATSRAGPLPPSPMPAVA 229
 DB 134 CPAGAPGPAPG-----PGA-----PGPKNNQPGAPAPAS-GARP-PGPRGP-A 175
 QY 230 GAAGGJLALLLGVAGAGAMCWRRAKPSRHPG-PSFGRG--STGLGGGGM----- 283
 DB 176 GDAG-----SPQPGHPSGPGNPGSGRGLPSPSRPP 212
 QY 284 -GPREAPGELGIALRGGAADP-PFCPHYEKVSGDYGHVYIVODGPQSPPN 335
 DB 213 PGF-AGGPGPG--HSGAGSGPGPGP-----PGSGQPGHSGNDGVGAPGN 257

RESULT 19
 Q17208 PRELIMINARY; PRT; 569 AA.
 ID Q17208;
 AC Q17208;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Collagen (Fragment).
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Bombycoidea; Bombycidae; Bombyx.
 OX NCBI_TaxId=7091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BURROBAN 200X300; TISSUE=IMAGINAL WING DISC;
 RA MEDLINE=97148965; Pubmed=8995790;
 RT "Identification of a Bombyx collagenous protein with multiple short
 RT domains of Gly-Xaa-Yaa repeats: cDNA characterization and regulation

RT of expression".
 RL Insect Biochem. Mol. Biol. 26:677-685 (1996).
 DR EMBL; Z30348; CA83002.1; -
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR001419; Glutenn.
 DR Pfam; PF01391; Collagen; 5.
 DR PRINTS; PR00210; GLUTENIN.
 FT NON_TER 1
 SQ SEQUENCE 569 AA; 56583 MW; FE371482F0A3B0AF CRC64;
 Query Match 8.1%; Score 149; DB 5; Length 569;
 Best Local Similarity 29.3%; Pred. No. 0.0018;
 Matches 51; Conservative 14; Mismatches 77; Indels 32; Gaps 6;

QY 171 PRGAVPRKPVSEMPERDRGAHSLERKENTLPGDPTSNATSRAGPLPPSPMPAVA 230
 DB 1 PGQGGQPIKPGQ-----PGYPGPGQPGYGPQ-GQPGAPGQPGQPGQPGQPG 50
 QY 231 AAGGJLALLLGVAGAGAMCWRRAKPSRHPG-----PGSPRGSLGGGGGM 283
 DB 51 QAG-----QPTPGQG-----QPIKPAQPGHPPGQPGQPGQPGQPGQPG 99
 QY 284 GPREAPGELGIALRG-----GAADPFCPHYEKVSGDYGHVYIVODGPQSP 335
 DB 100 PGQPGQPGQPGQPGQPGQPGQPGQPGQPGQPGQPGQPGQPGQPGQPG 153

RESULT 20
 Q19111 PRELIMINARY; PRT; 316 AA.
 ID Q19111;
 AC Q19111;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE F02D10.1 protein.
 GN F02D10.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Swinburne U.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; Pubmed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C. elegans*: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018 (1998).
 DR EMBL; Z67990; CA91932.1; -
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR002486; Col cuticle_N.
 DR Pfam; PF01391; Collagen; 2.
 DR Pfam; PF01484; Col cuticle_N; 1.
 DR Pfam; PF01484; Col cuticle_N; 1.
 SQ SEQUENCE 316 AA; 29444 MW; 4D76D5BA07923499 CRC64;
 Query Match 8.0%; Score 148.5; DB 5; Length 316;
 Best Local Similarity 28.5%; Pred. No. 0.001;
 Matches 49; Conservative 12; Mismatches 78; Indels 33; Gaps 6;

QY 168 GQSPRGAVPRKPVSEMPERDRGAHSLERKENTLPGD-----TSNATSRGA 216
 DB 144 GASKGGAAPCEPSTPPCQ-PCPAGPGPGPGCTGEGGPGAPGAPGAPGAPGAPG 202
 QY 217 EGRLPPSPMPAVAGAGJLALLLGVAGAGAMCWRRAKPSRHPG-GSPRGGSL 275
 DB 203 PGPAGPGPNQPGQPGQPGG-ODGASSAG-----EAGPAGPGPAGPAGPAG 253
 QY 276 GLGGGGGKGP-----EAPGELGIALRGGAADPFCPHYEKVSG 316
 DB 254 GSGAGGPGPGPAGPAGPAGPAGSDGNPGTAGPAGNPGGEGEGKICPKYCAIDG 305

```

RESULT 21
ID Q61434 PRELIMINARY; PRT; 1140 AA.
AC Q61434;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Collagen (Fragment).
OS COL15A1.
SN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Abe N., Muragaki Y., Yoshitaka H., Inoue H., Nimomiya Y.;
RT "Identification of a novel collagen chain represented by extensive
RT interruptions in the triple-helical region."
RL Cell. Mol. Biol. Res. 196;576-582(1993).
DR EMBL; D17546; BA04483.1; -.
DR HSSP; P39061; IKOE.
DR MGD; MGI:88449; Col15a1.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR003129; TSPN.
DR Pfam; PF01391; Collagen; 8.
DR Pfam; PF02210; TSPN; 1.
DR ProDom; PD000007; Collagen; 1.
FT NON_TER 1
SQ SEQUENCE 1140 AA; 115156 MW; 880C7B682B3BDFE CRC64;

Query Match 7.9%; Score 146.5; DB 11; Length 1140;
Best Local Similarity 23.7%; Pred. No. 0.0062;
Matches 92; Conservative 22; Mismatches 107; Indels 167; Gaps 21;

QY 2 GPPHSGPGVAVGALLIGVIG-LVSGLSLEPVYVNSANKRFQAGGYLYPQIGRDL 60
DB 164 GPP--GPAQPQ-----GPAQPVQSPNSQPV----- 187
QY 61 LCPRAR-PPGPPSPNVEFYKLYLVGAGQRC----- 93
DB 188 --PGAQGPFGQGPGRKX-----GTPGRDGEFGDGERGCDTPOGPGTDPDVG 237
QY 94 -----AAPANLLLTCDRPLDLRTIKQEVSPNLMGHEFRSHDYIIAT 140
DB 238 PKGKGDGPIGPRGPBP-----PGPGRBFRQDKLTIDME 274
QY 141 SDGTREGLESIQGVCLTRGAKVLRVQSGPRGAVPRKP--VSEMPEMRD-GAASLE 197
DB 275 GSGFGSDLESIRG-----PRGPPGPPGPGVGLPGEPRFGINSYA 317
QY 198 PKEKTLDPDPTSNATSRGAEGP--LP-PPSMPAVAGAGLALLLVAGAGAWCMRRR 254
DB 318 PGPALPLPVP-----GKEGPRGPPGPPGPRGKGP-----PGVAGKXSV----- 360
QY 255 RAKPESRHPGP-GSFGGSGIGCGGCGMGPRAREPELIGALRG--GAADPPCPHY 311
DB 361 ----GDVGI-PGPKSGKGLDIPGPKSGGLA---GSPGPVGPGRPPGPPGPGFAAGF 413
QY 312 EKVSQGYGHPVYVQ-----DGPQSP 333
DB 414 DDMBSG-GIPLWTTAKSSDGLQGPPGSP 440

```

```

OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
CX NCBI_TaxId=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=92011898; PubMed=1918153;
RA Su M.W., Suzuki H.R., Bleker J.C., Solursh M., Ramirez F.;
RT "Expression of two nonallelic type II procollagen genes during
RT laevis embryogenesis is characterized by stage-specific production of
RT alternatively spliced transcripts."
RL J. Cell Biol. 115:565-575(1991).
DR EMBL; M63586; AA49679.1; -.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000885; Fib_collagen_C.
DR InterPro; IPR001007; WVC_C.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 18.
DR Pfam; PF00093; WVC; 1.
DR ProDom; PD000007; Collagen; 3.
DR ProDom; PD002078; Fib_collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
DR SMART; SM00214; WVC; 1.
DR PROSITE; PS01208; WVC; 1.
KW Collagen.
SQ SEQUENCE 1491 AA; 142495 MW; 43026FF08FB0314 CRC64;

Query Match 7.9%; Score 146.5; DB 13; Length 1491;
Best Local Similarity 31.1%; Pred. No. 0.0085;
Matches 56; Conservative 11; Mismatches 80; Indels 33; Gaps 8;

QY 171 PGCAVPRKPEVSEMPMERDRG-AAHSLPEKENTLPDPTSNATSRGAGPLPPSMPAVA 229
DB 123 PKGPPGPGPSGEGRGRGDXGTGA-PGPRGRGP-----GTRGNQGPAGPPGP 178
QY 230 GAAGTALLILIG--VAGAGAMCMRR-----RAKPSBRHGPQSF--GRGSLGLG 278
DB 179 GIGGPMFAAMGTGFEKKGAGCMGVQGMGPRGPGPGGAGGPGGQGNPEPES 238
QY 279 GGGG-WGPR-----EAPELIGALRG-----GAADPPCPHYEVSQGYDP 321
DB 239 GAGGPMGPRGPGPSGKEDDSEAKGKSGRGPFGQAGRGFTGTGLPQVKHRTY 298

```

```

RESULT 23
ID Q62001 PRELIMINARY; PRT; 1774 AA.
AC Q62001; Q60672;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Procollagen, type XVIII, alpha 1 precursor (XVIII) collagen
DE (Procollagen, type XVIII, alpha 1) (Alpha-1 type XVIII collagen).
GN COL18A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=PRIMARY TAIL CULTURE;
RX MEDLINE=94245707; PubMed=8186673;
RA Rehn M., Hintikka E., Pihlajaniemi T.;
RT "Primary structure of the alpha 1 chain of mouse type XVIII collagen,
RT partial structure of the corresponding gene, and comparison of the
RT alpha 1 (XVIII) chain with its homologue, the alpha 1(XV) collagen
RT chain."
RL J. Biol. Chem. 269:13929-13935(1994).
RN [2]
RP SEQUENCE OF 1-562 FROM N.A.
RX MEDLINE=94240112; PubMed=8183894;
RA Rehn M., Pihlajaniemi T.;

```

RT "Alpha 1 (XVIII), a collagen chain with frequent interruptions in the
RT collagenous sequence, a distinct tissue distribution, and homology
RT with type XV collagen.",
RN Proc. Natl. Acad. Sci. U.S.A. 91:4234-4238 (1994).
RN [3]
RP SEQUENCE OF 1-562 FROM N.A.
RX MEDLINE=95181468; PubMed=7876242;
RA Rehn M., Philajantien T.,
RT "Identification of three N-terminal ends of type XVIII collagen chains
RT and tissue-specific differences in the expression of the corresponding
RT transcripts. The longest form contains a novel motif homologous to rat
RT and Drosophila fibrized proteins.",
RL J. Biol. Chem. 270:4705-4711 (1995).
DR EMBL, U03715; AAC52903.1; -
DR EMBL, U03716; AAC52903.1; JOINED.
DR EMBL, U03718; AAC52903.1; JOINED.
DR EMBL, U03718; AAC52903.1; JOINED.
DR EMBL, U34607; AAC52903.1; JOINED.
DR EMBL, U34608; AAC52903.1; JOINED.
DR EMBL, U34609; AAC52903.1; JOINED.
DR EMBL, U34610; AAC52903.1; JOINED.
DR EMBL, U34611; AAC52903.1; JOINED.
DR EMBL, U34612; AAC52903.1; JOINED.
DR EMBL, U34613; AAC52903.1; JOINED.
DR EMBL, U11637; AAC52179.1; -
DR HSSP, P39061; 1KOE.
DR MGD, MGI:88451; Col18a1.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR001791; Lamlnh_G.
DR InterPro; IPR003129; TSPN.
DR Pfam; PF01391; Collagen; 8.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF02210; TSPN; 1.
DR ProDom; PD000007; Collagen; 1.
DR SMART; SM00063; FRI; 1.
DR SMART; SM00282; Lamc; 1.
DR SMART; SM00210; TSPN; 1.
DR PROSITE; PS00319; Fz; 1.
DR Signal.
SQ SEQUENCE 1774 AA; 182229 MW; CF4D9BC988E232 CRC64;
Query Match 7.9%; Score 146.5; DB 11; Length 1774;
Best Local Similarity 23.7%; Pred. No. 0.01;
Matches 92; Conservative 22; Mismatches 107; Indels 167; Gaps 21;
QY 2 GPPHSGPGGVAVGALLLLGLVGLVSGLSLEPYVWNSANKRFQAGSGYLYPQIGDRLDL 60
DB 798 GPP-GRAPGQ-----GRAPGVVGSNSQPV----- 821
QY 61 LCPRAA-PPGPHSSPNYEFYKLYLVGAQGRCE----- 93
DB 822 --PGAQGPFGPGPGPKD-----GTPGRDGPDPGPDGPGGTGPGVDVG 871
QY 94 -----APPAPYLLLTCDRLFTIKFGYSFNLWGHESHHDYIAT 140
DB 872 PKGEKADPGIGRPGPGP-----DPPGPPSRQKLTFFIDME 908
QY 141 SPGTREGLSLQGVCLTRGMKYLNRVGSFRGAVFRKP--VSEMPMERDR-GAASLE 197
DB 909 GSGFSGDIESLRG-----PRGFPGPGPGPGVCGIPGPRGFRFGNGSYA 951
QY 198 PGKENLPGDPTSNATSRGAGP--LP-PPSPYAVVGAAGLLLLLVGAAGAGMCMRR 254
DB 952 PGPAGLPVP-----GKGGPGPGPGPGPPPKGSGP-----PGVAGKGSV----- 994
QY 255 RAKPSRSRHPG-PSPGRGSLGLGGGGMGRBAPGELIALRG--GAADPPPCPHY 311
DB 995 ---GVVGPGRGSKGDIQPIGMKSGLA---GSPGVGPGPGPGPGPGPGFAAGF 1047
QY 312 EKSGDYGHVYIVQ-----DPPPSG 333
DB 1048 DMEGS-GIPLMTTARSDDLQGPSP 1074

RESULT 24
Q90800 ID Q90800 PRELIMINARY; PRT; 675 AA.
AC Q90800;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Collagen-alpha-3 type IX precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CARTRILAGE;
RX MEDLINE=92241276; PubMed=1572350;
RA Brewton R.G., Ouspenskaya M.V., van der Rest M., Mayne R.;
RT "Cloning of the chicken alpha 3(IX) collagen chain completes the
RT primary structure of type IX collagen.",
RL Eur. J. Biochem. 205:443-449 (1992).
DR EMBL, X64712; CAA45967.1; -
DR InterPro; IPR000087; Collagen.
DR Pfam; PF01391; Collagen; 11.
DR ProDom; PD000007; Collagen; 3.
KM Collagen; signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 675 COLLAGEN-ALPHA-3 TYPE IX.
SQ SEQUENCE 675 AA; 63069 MW; 95094637A39D72D CRC64;
Query Match 7.9%; Score 146; DB 13; Length 675;
Best Local Similarity 28.8%; Pred. No. 0.0037;
Matches 64; Conservative 9; Mismatches 71; Indels 78; Gaps 12;
QY 166 RVGSPRGAVPRKPVSEMPMERDRGAHSLPEKKNLPDPPSNATSRGABPPLPPSM 225
DB 23 RVG--PQPPPGPGPGPGSGKDIDG---EPGSGLPGP-----GPKAPGK 65
QY 226 PAVVGAAG--GLALLLVGAAGAGMCMRRRAKPSRSRHPG-----DPSFRG- 272
DB 66 PANGAGELPGLP---GVDTLTGT-----DPPGPNPGPRGALGPAGPGPAKGL 115
QY 273 -----GSLGLGGGGM---GPR-----EABPGLIALRG- 300
DB 116 GPPPGPGPGSLPGANGFRPPPGSLPGRPPGPGPGGLAGIIPGGGDIQCPALCP 175
QY 301 -GAADPPPCPHYEVNSGDYGHVYIVODG-----PQSPNI 336
DB 176 GPPPGPMGFGKGTGKGPGEIGKEKSGSPGPPPGI 217
RESULT 25
Q90ZS0 ID Q90ZS0 PRELIMINARY; PRT; 1669 AA.
AC Q90ZS0;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Alpha 3 collagen IV.
GN COL4A3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=20005934; PubMed=10534397;
RA Lu W., Phillips C.L., Killen P.D., Hlaing T., Harrison W.R.,
RA Elder F.F.B., Miner J.H., Overbeek P.A., Meisler M.H.;
RT "Insertional mutation of the collagen genes col4a3 and col4a4 in a
RT mouse model of alport syndrome.",

RL Genomics 61.113-124(1999).
 DR EMBL; AF169387; AAD50449.1; -
 DR MGI; MGI:104688; Col1a3.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR001442; Procollagnc4.
 DR InterPro; IPR000504; Rna_rec_mot.
 DR Pfam; PF01413; C4; 2.
 DR Pfam; PF01391; Collagen; 21.
 DR ProDom; PD000007; Collagen; 6.
 DR ProDom; PD003923; Procollagnc4; 1.
 DR SMART; SMO0111; C4; 2.
 DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
 KW Collagen.
 SQ SEQUENCE 1669 AA; 161769 MW; 30976E59739A47B2 CRC64;
 Query Match 7.9%; Score 146; DB 11; Length 1669;
 Best Local Similarity 24.3%; Pred. No. 0.011;
 Matches 97; Conservative 34; Mismatches 134; Indels 134; Gaps 22;
 QY 4 PHSQGGVAVALLLLGLGL-----VSGLSLEP-----VYNSANKRFGAEGGY 48
 DB 1008 PHMGGS-----MGIMVPGPKGRKGTSLPLAGRPGRGTGHGPQDK--GEPGY 1056
 QY 49 VLTPQIGRLDLCPRAPPGPHSSPNYEFYKLYLVGAQGRCE----- 93
 DB 1057 SEGARP-----PPGPKDPC-----LPDQKCKGRREVPPGPGSGPACP 1097
 QY 94 -AAPAPVLLITCDRP--DLDRFTTKFOEYSPNIMGHEFRSHNDYIIATSDGTREG 147
 DB 1098 DGAPSPSPSPGPHGPKFPAGDLG---KQCKGFPGPPG-----STGPPGPPG 1141
 QY 148 LESLQGGVCL-----TRGKVLRLVQSPRGGAVERKRVSEMPMERDRGA 123
 DB 1142 LPLGPGPKMGMDGGRDIPPPGKGTGLGAYPGPKS--PGVGA---KQDRGV- 1194
 QY 194 HSLBEGKENTLPDPTSNATSGAGCPLEPPSPMAVAGALALLLVAGAGGAMCWR 253
 DB 1195 ---PGLSGLPG-----RKGVMDVPGQPPGAPGPPGALIPGPKDRDLPG 1243
 QY 254 RRAVPSERHNG-PGSPGRC-----GSLGLGG-GGMPRAAEVGEGLALRG 300
 DB 1244 LRNGPGEPPGPPGPIGKIGDKGFNGPPGKGLPPTYGDMGP-DGFPAPGTPGLPG 1302
 QY 301 GAADPPCFHYEKVSGDYHPVY---IYQDGP--POSPP 334
 DB 1303 VRGDPGF-PGFPPIKGEKGNPGLPIGHPPGVPGKPP 1340

RESULT 26
 ID 017805 PRELIMINARY; PRT; 305 AA.
 AC 017805;
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE F14F7.1 protein.
 GN F14F7.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 CX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McMurtry A.A.;
 RL Submitted (NCV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 KW MEDLINE=99069613; PubMed=9851916;
 FX none;
 RA "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology.";
 RT Science 282:2012-2018(1998).
 RL EMBL; Z81503; CAB04111.1; -.

DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR002486; Col_cuticle_N.
 DR Pfam; PF01484; Col_cuticle_N; 1.
 SQ SEQUENCE 305 AA; 29497 MW; 161BFA90D024C2B6 CRC64;
 Query Match 7.9%; Score 145.5; DB 5; Length 305;
 Best Local Similarity 22.5%; Pred. No. 0.0016;
 Matches 83; Conservative 21; Mismatches 122; Indels 143; Gaps 12;

QY 11 VRVGLALLLG-----VLGLVSGLSLEPVYWN-----SANK----- 40
 DB 5 VALKAYRLAFSAVAFCLSVSVSCITLPMYNYNSGRVLYGVSVSPCKKSANAEVFEV 64
 QY 41 -----RQABGGVLYTPQIGRLDLCPRAPPPGPHSSPNYEFYKLYLVG 86
 DB 65 NHIRASATNSTRSAHAGYGY-----AQPQGGGG-----GG 97
 QY 87 AGGRCEAPPAPVLLITCDRLDRLFTTKFOEYSPNIMGHEFRSHNDYIIATSDGTRE 146
 DB 98 GQSTCCRRGPPGAGGTGKG-----RFGAPGAG----- 128
 QY 147 GLESLQGGVCLTRGKVLRLVQSPRGGAVERKRVSEMPMERDRGAHSLBEGKENTLPD 206
 DB 129 -----MPGNPGKGGSGPCHPVYPTPCPKCPG--RPGPPGPPGR 165
 QY 207 PHSN-----ATSRGA-----EGFLPPSPMAVAGALALLLVAGAGGAMCW 251
 DB 166 PGSDQPGRPATGGAAPPBPBPBPBPBPBPBPBPBPBPBPBPBPBPBPBPBPBPBP 225
 QY 252 RRRAPKPSERHNGPSPGSGSLGLGGGGV---GPREAPGELIAGGGAADPPF 307
 DB 226 AGPRGAPGAGHPPSGSGGGRPPGAPKAPQGPGRG-DGHPQPPGPPGSGSGGNRV 284
 QY 308 CPHYEKVSG 316
 DB 285 CPXYCAIDG 293

RESULT 27
 ID 094620 PRELIMINARY; PRT; 308 AA.
 AC 094620;
 DT 01-FEB-1997 (Tremblrel. 02, Created)
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Cuticle preprocollagen.
 GN COL-2.
 OS Meloidogyne incognita (southern root-knot nematode).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
 OC Tylenchoidea; Heteroderidae; Meloidogyinae; Meloidogyne.
 CX NCBI_TaxID=6306;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cuticle;
 RA MEDLINE=98324411; PubMed=9662034;
 RX Wang T., Deom C.M., Husey R.S.;
 RT "Identification of a Meloidogyne incognita cuticle collagen gene and
 RT characterization of the developmental expression of three collagen
 RT genes in parasitic stages.";
 RT Mol. Biochem. Parasitol. 93:131-134(1998).
 DR EMBL; U68729; AAC48358.1;
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR002486; Col_cuticle_N.
 DR Pfam; PF01391; Collagen; 2.
 DR Pfam; PF01484; Col_cuticle_N; 1.
 KW Collagen.
 FT CHAIN
 SQ SEQUENCE 308 AA; 29361 MW; 825AER0249PFC78 CRC64;

Query Match 7.9%; Score 145.5; DB 5; Length 308;
 Best Local Similarity 28.4%; Pred. No. 0.0017;
 Matches 46; Conservative 14; Mismatches 67; Indels 35; Gaps 6;

```

QY 171 PRGAVPRKRVSEMPERDGAHSLPEPKENLPDPTSNATSRGAEPLPPSPPAVAG 230
DB 155 PCRGGGPGGPGPPGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPG 204
QY 231 AAGGALLLLGAVAGAGAMCWRRRRAKPPESRHPG---PGSPFGGSGIGL-GGGGGNGP 285
DB 205 QPGG-----PGQGAQA-OSQEGAPAPPGPPGAGAPGQPGVNGPAGQPGGPGGPGP 255
QY 286 R-----EABPELGIALRGGAADPPPCPHYKVSQ 316
DB 256 KGPFGPGGPGNDGSPGQPGPPPTNGSGGKGCPCATCAIDG 297

RESULT 28
Q9ND02 PRELIMINARY; PRT; 680 AA.
AC Q9ND02;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Procollagen, type IX, alpha 1.
GN COL9A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO.
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai U., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Fukunishi Y., Komori H., Adachi U., Fukuda S.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann M., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuenli P., Lewis S., Matsumoto Y., Nakai T., Pesole G., Quackenbush J.,
RA Schmitt L.M., Stabili F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima U., Mazzarelli U., Mombereis P.,
RA Norone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang X.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszewski A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK011547; BAB27690.1; -.
DR MGD; MGI:88465; Col9a1.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF01391; COL1F; 1.
DR PRINTS; PR01582; KVJ3CHANNEL.
DR ProDom; PD000007; Collagen; 3.
KM KX
SQ SEQUENCE 680 AA; 64733 MW; 5C0FB3E992B4D91 CRC64;

Query Match 7.8%; Score 145; DB 11; Length 680;
Best Local Similarity 24.0%; Pred. No. 0.0045;
Matches 87; Conservative 22; Mismatches 133; Indels 122; Gaps 17;

```

```

DB 200 -----PRGQHKKEBPDQGLGEVGDQPPGQGLRGTGTGVGDKKKGARGFDEGPGQ 254
QY 175 AVP-----RKVSEMPERDGAHSLPEPKENLPDPTSNATSRGAEPLPPSP 224
DB 255 GIPGAGDQGGPGPPGTGEGDRGI-----QSGRGI-PGSPGKGDGLGVGDRDIPG 309
QY 225 MPVAGAAG-----GLALL--LLGAVAGAGAMCWRRRRAKPPSE-----SRHPG--- 265
DB 310 MPGRKGAGKRGPPRGDVLGGLPGVPGIPGAKVYAGKNTGAPGKRGQUGSSGKPGQG 369
QY 266 -PGSFGGSGIGL-GGGGGNGPRAE--PGLGIALRG--GAADPPCPHYKVSQDYG 319
DB 370 PPGVGPGRGRLPGSGRGVPGEGSPGIPGKLGSGVSGPLGLPGPGLPGMKGRGVGF 429
QY 320 HP 321
DB 430 EP 431

```

```

RESULT 29
Q9NZ06 PRELIMINARY; PRT; 1745 AA.
AC Q9NZ06;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Collagen type V alpha 3 chain.
GN COL5A3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA; AND HEART;
RX MEDLINE=20187594; PubMed=10722718;
RA Imamura Y., Scott I.C., Greenspan D.S.;
RT "The pro-alpha3 (V) collagen chain. Complete primary structure,
RT expression domains in adult and developing tissues, and comparison to
RT the structures and expression domains of the other types V and XI
RT procollagen chains."
RL J. Biol. Chem. 275:8749-8759(2000).
DR EMBL; AF177941; AAP53902.1; -.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000885; Fib_collagen_C.
DR InterPro; IPR001791; Laminin_G.
DR Pfam; PF01410; COL1F; 1.
DR Pfam; PF01391; Collagen; 17.
DR Pfam; PF02210; TSPN; 1.
DR ProDom; PD000007; Collagen; 3.
DR ProDom; PD002078; Fib_collagen_C; 1.
DR SMART; SM00038; COL1F; 1.
DR SMART; SM00282; LamG; 1.
DR SMART; SM00210; TSPN; 1.
KM KX
SQ SEQUENCE 1745 AA; 172051 MW; SE8FF97135397AC1 CRC64;

Query Match 7.8%; Score 145; DB 4; Length 1745;
Best Local Similarity 24.0%; Pred. No. 0.013;
Matches 88; Conservative 24; Mismatches 103; Indels 152; Gaps 20;

```

```

OY 174 GAVPRKPVSEMERDRCGAHSLPEKENTLPGDPTSNATSRGAGPLP-----PSMFA 227
DB 960 EUGPFGPGKEGEPAGLRG-----PGFPGGPGDPGPTGL-KGDKPGPPGVANGSPGERG 1013
OY 228 VAGAGGLALLLLGVAGAGGACWRRRAKPSSESHPG-PSFGGSLG-LGGGGGNGP 285
DB 1014 PLGPRAGGIC-LPGSGSGSPGVGPAKKGSGRGERPPPTGKDGIPGLGFLGPPGAAGP 1071
OY 286 --REAPGELGIALRNG-----GAADPPCPHYEKSGDYGHPYIVODGPP----- 330
DB 1072 SGEEDDKGVGAPGHKSGKXGKDAGPPGQP---GIRGPAHGHP-----GPPADGAQGR 1122
OY 331 GSPNNY 337
DB 1123 RGPPLGF 1129

RESULT 30
Q25466 PRELIMINARY; PRT; 309 AA.
AC Q25466;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE COL-1.
GN COL-1.
OS Meloidogyme incognita (southern root-knot nematode).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidea; Heteroderidae; Meloidogyminae; Meloidogyne.
OX NCBI_TaxID=6306;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97164124; PubMed=9010847;
RA Ray C., Wang T.Y., Husey R.S.;
RT "Identification and characterization of the Meloidogyme incognita coll
RT cuticle collagen gene.";
RL Mol. Biochem. Parasitol. 83:121-124(1996).
DR EMBL; U40766; AAC47437.1; -.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR002486; Col_cuticle_N.
DR Pfam; PF01391; Collagen_2.
DR Pfam; PF01484; Col_cuticle_N; 1.
SQ SEQUENCE 309 AA; 29472 MW; 0C6C9D5BA76C8916 CRC64;

Query Match 7.8%; Score 144.5; DB 5; Length 309;
Best Local Similarity 28.4%; Pred. No. 0.002;
Matches 46; Conservative 14; Mismatches 67; Indels 35; Gaps 6;

OY 171 PRGGAVERKPVSEMERDRCGAHSLPEKENTLPGDPTSNATSRGAGPLPSPMPAVAG 230
DB 156 PCPGGQPGPPGPPGPGGPGGPGGQAGSPGEP-----GPAGPPGSPGAPG 205
OY 231 AAGGLALLLLGVAGAGGACWRRRAKPSSESHPG-PSFGGSLG-LGGGGGNGP 285
DB 206 QGCG-----PGPGGSA-QSGAGAPAPPPGPDADAPGAPGAPGAGGPGGPGGPG 256
OY 286 R-----EAPGELGIALRNGGADPPFCPHYEKVSG 316
DB 257 KGPFGPPGQPGNDGAPGQGTGTPGAGKEGICPKYCALDG 298

RESULT 31
Q960B3 PRELIMINARY; PRT; 1347 AA.
AC Q960B3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-UTN-2002 (TREMBLrel. 21, Last annotation update)
DE Type V preprocollagen alpha 2 chain (Fragment).
GN COL5A2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21451029; PubMed=11566270;
RA Valkila M., Melkonian M., Kvist L., Kuivaniemi H., Tromp G.,
RA Ala-Kokko L.;
RT "Genomic organization of the human COL3A1 and COL5A2 genes: COL5A2 has
RT evolved differently than the other minor fibrillar collagen genes.";
RL Matrix Biol. 20:357-366(2001).
DR EMBL; AY016295; AAL13166.1; -.
DR EMBL; AY016289; AAL13166.1; JOINED.
DR EMBL; AY016290; AAL13166.1; JOINED.
DR EMBL; AY016291; AAL13166.1; JOINED.
DR EMBL; AY016292; AAL13166.1; JOINED.
DR EMBL; AY016293; AAL13166.1; JOINED.
DR EMBL; AY016294; AAL13166.1; JOINED.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000885; Fib_collagen_C.
DR Pfam; PF01410; COLFi. 1.
DR Pfam; PF01391; Collagen_18.
DR ProDom; PD000007; Collagen; 6.
DR ProDom; PD002078; Fib_collagen_C; 1.
KW Collagen.
FT NON TER.
SQ SEQUENCE 1347 AA; 128822 MW; F25F20E6B23A17C CRC64;

Query Match 7.8%; Score 144.5; DB 4; Length 1347;
Best Local Similarity 33.0%; Pred. No. 0.011; 77; Indels 35; Gaps 11;
Matches 61; Conservative 12; Mismatches 77; Indels 35; Gaps 11;

OY 169 QSPRG--GAV-PRKPVSEMERDRCGAHSLPEKENTLPG-----DPTSNATSRGAE 217
DB 351 RGRGRGPGTVGPPGPGVGEAGPGRNGF-----FGSDGLPFGPAGAGERPPVSSSGKSGQ 405
OY 218 GRLPPSPMAVAGAGGLALLLLGVAGAGGACWRRRAKPSSESHPG-PSFG-RG--G 273
DB 406 GDGRGPGEPGLPGAR-----LTGNFVGVPGRGKLPGLAPGEDGPFPPGSGIRGPGG 460
OY 274 SLGLGGGGGWPGEAPRGSLGIA-----LRGGADPPFCPHYEKSGDYGHPYIVODGP 329
DB 461 SMGLPPKXSSGDPGKRGAGNAGVGGGAGAPKQGEVGP-----SGPVPGGLAGERG- 514
OY 330 PGSP 334
DB 515 EQGPP 519

RESULT 32
Q90W37 PRELIMINARY; PRT; 1420 AA.
AC Q90W37;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-UTN-2002 (TREMBLrel. 21, Last annotation update)
DE Alpha 1 type IIA collagen precursor.
GN COL2A1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=STERNAL;
RA Caixia X., Yongzhi X., Siqi G., Yiyang S.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY046949; AAK96621.1; -.
DR InterPro; IPR000085; Fib_collagen.
DR InterPro; IPR000885; Fib_collagen_C.
DR Pfam; PF01410; COLFi. 1.
DR Pfam; PF01391; Collagen; 18.
DR ProDom; PD000007; Collagen; 6.
DR ProDom; PD002078; Fib_collagen_C; 1.

```

KM Collagen: Signal.
 FT SIGNAL 1 25 POTENTIAL.
 SQ SEQUENCE 1420 AA; 13499 MW; 88D9AAB17F214FF5 CRC64;
 Query Match
 Best Local Similarity 28.1%; Score 144.5; DB 13; Length 1420;
 Matches 65; Conservative 11; Mismatches 90; Indels 65; Gaps 10;

QY 147 GLESLQG--GVCITRGMKTLIRVQSGFRGAVVRKP-----VSEMER----- 188
 DB 644 GAQGLGQPRGLPGTPTG-----DGPKGATGPAAGNAGQPPGLQGMGKERGAAGIAGL 696
 QY 189 --DRGAHSLPEPKENLPDPTSNATSRGAEPLPPSPMAVAGAGALLLGVAGAG 246
 DB 697 KGRGVDVG--EKPEGAFG---KDGAKLTGPIGPPGAPGVEKESG--PPGSGAA 748
 QY 247 GANCMRRRAKPPESRHPG--PGSFRGSGSLGIGGGGMPREAPBELGIALRG----- 300
 DB 749 GA-----RGAPGERGPGAPGAPGAPGAGDQPGAKGEQEPQKDGADAGAPQGPS 802
 QY 301 -----GAADPPGCHYEKVSQDYGHPYIVQDPPPS 333
 DB 803 GADPGQPGTGTGPKAGARGAQGPATGFPAGAGVGPQPGNPGPP 853

RESULT 33
 Q9F342 PRELIMINARY; PRT; 775 AA.
 AC Q9F342;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
 DE Hypothetical protein SC05128.
 GN SC05128 OR SC9812.13.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Lake L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorek A., Woodward J., Barrrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL: AL391751; CAC05758.1; -.
 KM Hypothetical protein.
 SQ SEQUENCE 775 AA; 75820 MW; E5032658A20A0711 CRC64;

Query Match
 Best Local Similarity 7.8%; Score 143.5; DB 16; Length 775;
 Matches 94; Conservative 23; Mismatches 130; Indels 153; Gaps 19;

QY 3 PHSGPG-----GVRVGLLILGLVLSGLSLPEPYWNSANKRFOAEGGYL 50
 DB 133 FRSGSGPKTQGGSGTGAAGLIPAGSAPTAGTPAAGAGAGAGAGAG----- 181
 QY 51 YPOIGRLDLCPARPPGHSSPNVEFYLYLVGAQO-----GRCFA---PPAPNLL 101
 DB 182 -----ARGGSGRPGGVGSMAPGAGRSSAGGARSATGTPVAP--- 224
 QY 102 LTCRPLDLRFTTKFOEYSPNLMGHEFRSHHDYIIATSDGTREGLESLOGVCLTRGM 161
 DB 225 -----GH-----GGGTGSGFDVTEALAGPILGNGN 249

QY 162 KYLIRVQSGSP-----GAIVPRKPVSEMPMERDRGAHSLPEPKENLPDPTSNATSRGAE 217
 DB 250 GNGNGNSGSPFPDADGEBPR--DDLPTFAGEBRAAOAPGQGN--GLNGQSYGTGPQ 305
 QY 218 GPLPPSPMAV-----AGAGGLALLLGVAGAGANCMWRRAKPPESRHP----- 264
 DB 306 GPAPBTGSPAPGDSRLTPPAGDGLG-----LGTAGCPAA-----POTQSEAPANG 351
 QY 265 --GPGSF--GRCGSLGIGGGGMPREAPBELGIALRGGAAD-----PPCPH 310
 DB 352 LRGPISLGGPSPGPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAP 405
 QY 311 YEKVSD---YGHFPIVQ-----DGP--QSPENI 336
 DB 406 PDNISGNTVTSIGIPVPEERTAPPAPGAGSDGPRPHTPKL 445

RESULT 34
 Q26640 PRELIMINARY; PRT; 1747 AA.
 AC Q26640;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
 DE Alpha2 (IV)-like collagen.
 GN COLP4ALPHA.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotus.
 OX NCBI_TaxID=7668;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94230414; PubMed=8175744;
 RA Expósito J.Y., Suzuki H., Geourjon C., Garrone R., Solursh M.,
 RA Ramirez F.;
 RT "Identification of a cell lineage-specific gene coding for a sea
 urchin alpha.2(IV)-like collagen chain.";
 RL J. Biol. Chem. 269:13167-13171(1994).
 DR EMBL: X76730; CAA54146.1; -.
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR001442; ProcollagenC4.
 DR Pfam: PF01413; C4; 2.
 DR Pfam: PF01391; Collagen; 21.
 DR ProDom: PD000007; Collagen; 6.
 DR ProDom: PD003923; ProcollagenC4; 1.
 DR SMART: SM00111; C4; 2.
 KM Collagen.
 SQ SEQUENCE 1747 AA; 173312 MW; EE722E878394B986 CRC64;

Query Match
 Best Local Similarity 7.7%; Score 143; DB 5; Length 1747;
 Matches 84; Conservative 24; Mismatches 159; Indels 82; Gaps 14;

QY 7 GPGYRVGALLLGLVLSGLSLPEPYWNSANKRFOAEGGYLYPIQIGRLDLCPRAR 66
 DB 266 GPRGMDG---MKATGEVGLDLSGYDYGKGLPGYSGERGFPENPEIGLLNGNGEKR 321
 QY 67 PPGHSSPNVEFYLYLVGAQGRCEAPAPNLLLCRPPDLRFTIKFOEYSPYLMWG 126
 DB 322 -DGQGRGNYGK--GPGSDYGMDDPGRAPDL----- 352
 QY 127 HEFRSHHDYIIATSDGTREGLESLOGVCLTRGMKYLIRVQO---SPRGAVPRKPV 182
 DB 353 -----EELVGVGPDPDEGPPGNPGRGSGAIGLFGQGGPYGPPMGPPRPTG 405
 QY 183 EMPMERDRGAHSLPEPKENLPDGP-----TSNATSRGAGPL-----PPSPMAVAGAA 232
 DB 406 SGGRSDDEKESGRPGIPGQSETGRANDGNRGEKGMWESRPGFMDGSTRQ 465
 QY 233 GGLALLLGLVAGAGANCMWRRAKPPESRHPG-----PGSFRGSGSLGIGGGGMP 285
 DB 466 G-----FMQKGRGSP-----PGRAGPAGPARSGNQSFQGFPGDTGNTGLKGMGIGALG 516

QY 286 REAFEGELIALRGGAADPPCPHYEK-VSGDGHPPYIVQDGPQSP 333
 DB 517 RDGRGSKG---ELGIGC-PPCPGKGYPGDRGYP-----GDPGSP 554

RESULT 35

Q96D07 PRELIMINARY; PRT; 744 AA.

AC Q96D07;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 73.4 kDa protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Eukarya; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RA Strausberg R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC013581; AAH13581.1; -
 DR InterPro; IPR001073; Clq.
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF00386; Clq; 1.
 DR Pfam; PF01391; Collagen; 7.
 DR ProDom; PD000007; Collagen; 2.
 DR PROSITE; PS0113; Clq; UNKNOWN 1.
 KM Collagen; Hypothetical protein.
 SQ SEQUENCE 744 AA; 73364 MM; 2BC1B0955DE2C9A3 CRC64;

Query Match 7.7%; Score 142; DB 4; Length 744;
 Best Local Similarity 23.3%; Pred. No. 0.0084;
 Matches 72; Conservative 16; Mismatches 99; Indels 122; Gaps 13;

QY 67 PPGHSSPNTYFYKLYLVGAGRCAPAPNLLTCDRPLDLRFTIKGEVSPILMG 126
 DB 198 PPGHGLP-----IKPGGPGLPQGPQPK----- 222

QY 127 HFRSHHDYIATSDGTREGLSELOGGVCLTRGMKVLRLVQSPRGAVPRKPVSEPM 186
 DB 223 -----GDRGKGLPGPQ-----LRGPKGKGCWPGAVKQKP- 256

QY 187 ESDRGAASLEPKENLPDPTSNATN-RGAEGLPPPSMAVAVAGALALLL---G 241
 DB 257 ---PQMG-PGPGVGLPGVGKPGVTGPPGPGPLGKPGAPGPGPGIGVQGP 311

QY 242 VAGAGCWCRRRAKPSRHPG-----PGSFGRGGS 274
 DB 312 IIGIG-----KPGDDGIPGPGPGPGGKGGGLPGLPQPLPGIGKPGPGPGD 361

QY 275 LGLGG-GGGMGRREAPGELGIALRGGAADPPCFHYEKVSGDYGHVYIVQDGP 329
 DB 362 RGMGVPGLGPR-GEKPGIGAPGIGG---PPGPGPLGPGWPGPGALGPGPKGEG 416

QY 330 ---PQSP 334
 DB 417 GIVPGQSP 425

RESULT 36

Q91A04 PRELIMINARY; PRT; 1835 AA.

AC Q91A04;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Alpha 1 (V) collagen.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.
 CX NCBI_TaxId=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=REED WHITE LEHOEN;
 RX MEDLINE=20068042; PubMed=10601735;
 RA Gordon M.K., Marchant J.K., Foley J.W., Igoe F., Gibney E.P.,
 RA Nah H.D., Barembaum M., Myers J.C., Rodriguez E., Dublet B.,
 RA van der Rest M., Linsmeayer T.F., Upholt W.B., Birk D.E.;
 RT "Complete primary structure of the chicken alpha1(V) collagen chain."
 RL Matrix Biol. 18:481-486(1999).
 DR EMBL; AF137273; AAF28099.1; -
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR000885; FIB_collagen_C.
 DR InterPro; IPR001791; Laminin G.
 DR InterPro; IPR001230; Premyl_site.
 DR InterPro; IPR003129; TSPN.
 DR Pfam; PF01410; COLFI; 1.
 DR Pfam; PF01391; Collagen; 19.
 DR Pfam; PF02210; TSPN; 1.
 DR ProDom; PD000007; Collagen; 2.
 DR ProDom; PD002078; FIB_collagen_C; 1.
 DR SMART; SM0038; COLF; 1.
 DR SMART; SM00282; LamG; 1.
 DR SMART; SM00210; TSPN; 1.
 DR PROSITE; PS00294; PRENYLATION; UNKNOWN 1.
 KM Collagen.
 SQ SEQUENCE 1835 AA; 184234 MM; D05B9D71022D44B2 CRC64;

Query Match 7.7%; Score 142; DB 13; Length 1835;
 Best Local Similarity 25.7%; Pred. No. 0.024;
 Matches 78; Conservative 15; Mismatches 82; Indels 128; Gaps 18;

QY 68 PPGHSSPNTYFYKLYLVGAGC-----RCGAPAPNLLTCDRPLDLRFTIKF 117
 DB 566 PPGMGPPG-----SGGKGEVEMGPGQPGRGIGGPPG-----AGP----- 602

QY 118 QEYSPMLGHEPSSHHDYIATSDGT-----AGSDGAKMPGQGTGRGGLAGLP 636
 DB 603 -----GRGR-----AGSDGAKMPGQGTGRGGLAGLP----- 636

QY 164 LRLVGSPPGGAVERKPVSEMERDRGAASLEPKENLPDPTSNATSRGAEGLP 223
 DB 637 ---EKGNRGEPGHPGPPGPGGDEGRG--DDGEVPRGLPQEP-----GPRLLGPKGP 686

QY 224 SMRAVAGAGLALLLGVAGAGCWCRRRAKPSRHPG-PGSFGGSLG-CCGG 281
 DB 687 GPPGPGVAG---WDGTPKGNV-----GPGEPGPPGQGNPGAGLP 732

QY 282 GMPRAEPGELGIALRG-----GADPPCFHYEKVSGDYGHVYIVQDGP 331
 DB 733 PIGP-----PGEKPRLLKPLPGWPGADGP-----GHP---GKEGPPGKXG 773

QY 332 SPP 334
 DB 774 GPP 776

RESULT 37

Q25582 PRELIMINARY; PRT; 284 AA.

AC Q25582;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Cuticular collagen.
 GN COLOST-1.
 OS Teladorsagia circumcincta.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdida; Strongylida;
 OC Trichostrongyloidea; Haemonchidae; Oestertaglinae; Teladorsagia.
 CX NCBI_TaxId=45464;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=97039674; PubMed=8885226;
 RA Johnstone I.L., Shafi Y., Mjaeed A., Barry J.D.;
 RT "Cuticular collagen genes from the parasitic nematode *Ostertagia circumcincta*.";
 RL Mol. Biochem. Parasitol. 80:103-112(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Johnstone I.L.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20017523; PubMed=10551361;
 RA Britton C., Redmond D.L., Knox D.P., McKerrrow J.H., Barry J.D.;
 RT "Identification of promoter elements of parasite nematode genes in *transgenic Caenorhabditis elegans*.";
 RL Mol. Biochem. Parasitol. 103:171-181(1999).
 DR EMBL; X96731; CAA6506.1; -;
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF01391; Collagen; 3.
 DR Pfam; PF01484; Col_cuticle_N; 1.
 DR Collagen.
 KM SEQUENCE 284 AA; 27534 MW; 60775B7D2CC40C77 CRC64;
 SQ

Query Match 7.6%; Score 141; DB 5; Length 284;
 Best local similarity 27.2%; Pred. No. 0.0033;
 Matches 61; Conservative 15; Mismatches 86; Indels 62; Gaps 12;

QY 121 SPNLGHEFRSHHDYIATSDTREGLESLOGGVCLTRGMKVLRYVQSPRGAVPRKP 180
 DB 62 SGNIRREVTIRQ---VLAAP--TRKARQS-GGGCC-----GGGVSPAGPPGP 102
 QY 181 VSEMPERDRGAHSLPEKENTPGDPTSNATSRGAGPLP-----PPSPAVAGAA 232
 DB 103 -----PGDGGPDSGDRGQGRNGPGRPATPAPDVAPCFNCPPGPPAGAP 151
 QY 233 GGLALLLGVAGAGACWRRRAKPSRHP-GPGSFRGSGSLGIGGGGGMGPRAEPG 291
 DB 152 GG-----RGGGSPGS-----DQPGNAGNPGGPGPIGPPGPPGNAQPG--NAGSPG 197
 QY 292 ELGILRGGAADPPFCPHYEKVSGDYGHVYIVODGPPGSPN 335
 DB 198 AFGTLTGPGSGQGPGRP-----GPGGP--GPDGQPGGPGN 232

RESULT 38
 Q25581
 ID Q25581 PRELIMINARY; PRT; 284 AA.
 AC Q25581;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Cuticular collagen.
 GN COL08T-2.
 OS Teladorsagia circumcincta.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongyloidea;
 OC Trichostrongylidae; Haemonchidae; Ostertagiinae; Teladorsagia.
 OC NCBI_TaxID=54564;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97039674; PubMed=8885226;
 RA Johnstone I.L., Shafi Y., Mjaeed A., Barry J.D.;
 RT "Cuticular collagen genes from the parasitic nematode *Ostertagia circumcincta*.";
 RL Mol. Biochem. Parasitol. 80:103-112(1996).
 DR EMBL; X96731; CAA6507.1; -;
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF01391; Collagen; 3.
 DR Pfam; PF01484; Col_cuticle_N; 1.
 DR Collagen.
 KM SEQUENCE 284 AA; 27560 MW; 7B7EBDD2DD41D76 CRC64;
 SQ

Query Match 7.6%; Score 141; DB 5; Length 284;
 Best local similarity 27.2%; Pred. No. 0.0033;
 Matches 61; Conservative 15; Mismatches 86; Indels 62; Gaps 12;

QY 121 SPNLGHEFRSHHDYIATSDTREGLESLOGGVCLTRGMKVLRYVQSPRGAVPRKP 180
 DB 62 SGNIRREVTIRQ---VLAAP--TRKARQS-GGGCC-----GGGVSPAGPPGP 102
 QY 181 VSEMPERDRGAHSLPEKENTPGDPTSNATSRGAGPLP-----PPSPAVAGAA 232
 DB 103 -----PGDGGPDSGDRGQGRNGPGRPATPAPDVAPCFNCPPGPPAGAP 151
 QY 233 GGLALLLGVAGAGACWRRRAKPSRHP-GPGSFRGSGSLGIGGGGGMGPRAEPG 291
 DB 152 GG-----RGGGSPGS-----DQPGNAGNPGGPGPIGPPGPPGNAQPG--NAGSPG 197
 QY 292 ELGILRGGAADPPFCPHYEKVSGDYGHVYIVODGPPGSPN 335
 DB 198 AFGTLTGPGSGQGPGRP-----GPGGP--GPDGQPGGPGN 232

RESULT 39
 Q17038
 ID Q17038 PRELIMINARY; PRT; 319 AA.
 AC Q17038;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 31.3 kDa protein.
 GN T15B7.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA "The sequence of *C. elegans* cosmid T15B7.";
 RT Investigating biology. The *C. elegans* Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Pauley A., Gattung S.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF022985; AAB69959.1; -;
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR002486; Col_cuticle_N.
 DR Pfam; PF01391; Collagen; 3.
 DR Pfam; PF01484; Col_cuticle_N; 1.
 DR Hypothetical protein.
 KM SEQUENCE 319 AA; 31292 MW; 5242AAEC68A488 CRC64;
 SQ

Query Match 7.6%; Score 141; DB 5; Length 319;
 Best local similarity 25.6%; Pred. No. 0.0036;
 Matches 90; Conservative 27; Mismatches 111; Indels 124; Gaps 23;

QY 13 VGALLLGVGLVLSGLS-----EPYVNSANKRFOAGSG---YVLPQIGDRDLIC 62
 DB 1 MSASTLVVAASAGIAIVCVFTVGMIFNDINSFYDEKIGELKFRGQIDA--WQAMI 58
 QY 63 PRAPPGRHSSPNVEFKLYVY-----GAQRRCBA--PRANLLLTCDRPPDL 111
 DB 59 PTPRPSGSS-----FLGNKQAEKNCGEQSGRGCPAGPPPP-----GQPG--- 101

```

QY 112 RTTFFQSYSPMLWGHFRSHHDYITATSDTREGLESLOGVCLTRGMKVLIRVQOSP 171
D 102 ---ARGAGLPLPAGQ-----PGSGARINPATGPPFCIT-----CP 135
QY 112 RGGAIVPRKPVSEMPERDRGAHSLPEPKENLPDPTSNATSRGAEGLPPSPMPAVAGA 231
D 136 AGAPPPAPG-----PGA-----PPKANNQGPAPAPAS-GGRP-PGRPP-AGD 177
QY 232 AGGALLLILVAGAGAGACWERRRRRAKPSRHPG-PGSFGRG--SLGLGGGGG---MG 284
D 178 AG-----SPQPGHPCSPGNCRGGRSRRTPGASGRPPGPG 214
QY 285 PREAPRGLALRGGAADP-PFCPHYKVSQDYGHPVYIVODGPPQSPPN 335
D 215 PAGA-PGQGG---RSGGAGTPOPGP-----PGSPGQPGHSGNDGVPTPKN 257

```

```

RESULT 40
QY 093485 PRELIMINARY; PRT; 809 AA.
AC 093485;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Alpha 1 type I collagen (Fragment).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Proactinopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FIBROBLAST;
RA Saito M., Kunisaki N., Hirono I., Aoki T., Ishida M., Urano N.,
RA Kimura S.,
RT Partial characterization of cDNA clones encoding the three distinct
RT pro alpha chains of type I collagen from rainbow trout.";
RL Fisheries Sci. 64:780-786(1998).
DR EMBL; AB008373; BAA3380.1; -.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000885; Fib_collagen_C.
DR Pfam; PF01410; COLFI; 1.
DR ProDom; PD000007; Collagen; 9.
DR ProDom; PD002078; Fib_collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
KW Collagen.
KW NON TER.
SQ SEQUENCE 809 AA; 78164 MW; 68C056A7640FC8A1 CRC64;

```

```

Query Match 7.6%; Score 141; DB 13; Length 809;
Best Local Similarity 27.4%; Pred. No. 0.011;
Matches 61; Conservative 15; Mismatches 93; Indels 54; Gaps 9;

```

```

QY 139 ATSDGTREGLESLOG-----GVCLTRGMKVLIRVQSPRGAVPRKPVSEMPERDRGA 193
D 98 ADGAGKKGCVRMGTPIPNPGAGSPGDK-----GETAPAVGPGSARGAPEGREGSGA 152
QY 194 HSLPEKKNLPDPTSNAT-----SRGABGP.PPPSPMPAVAGAGLALLL 239
D 153 ---PGPAGFAPPGGDDGPGAKGKAGDNGAKGDGAGGAPGFTGAPGPPGAPN----- 203
QY 240 LGVAAGAGACWERRRRRAKPSRHPG-----PSFGRGSLGLGGGGGMPREAPGE 292
D 204 TGAKKARGA-----AGPGATGFPGAAGRPPGSPGNNPPTGPGGKGGKGNRGE 257
QY 293 LGIALRGG--GAADPPFCPHYKVSQDYGHPVYIVODGPPQSP 333
D 258 TGPAGRPGLGAAGP-----GPXGKGGQPGGDPNPGSGTP 294

```

RESULT 41

```

Q910C0
ID Q910C0 PRELIMINARY; PRT; 1449 AA.
AC Q910C0;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Collagen a1(I).
GN COL1A1.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Proactinopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=1257802; PubMed=11358497;
RA Saito M., Takenouchi Y., Kunisaki N., Kimura S.;
RT "Complete primary structure of rainbow trout type I collagen
RT consisting of a1(I)a2(I)a3(I) heterotrimer.";
RL Eur. J. Biochem. 268:2817-2827(2001).
DR EMBL; AB052835; BAB5661.1; -.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000885; Fib_collagen_C.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 18.
DR ProDom; PD000007; Collagen; 2.
DR ProDom; PD002078; Fib_collagen_C; 1.
DR PROSITE; PS01208; WVFQ; UNKNOWN_1.
SQ SEQUENCE 1449 AA; 137117 MW; 62EBF8A7BFD652B8 CRC64;

```

```

Query Match 7.6%; Score 141; DB 13; Length 1449;
Best Local Similarity 27.4%; Pred. No. 0.022;
Matches 61; Conservative 15; Mismatches 93; Indels 54; Gaps 9;
QY 139 ATSDGTREGLESLOG-----GVCLTRGMKVLIRVQSPRGAVPRKPVSEMPERDRGA 193
D 738 ADGAGKKGCVRMGTPIPNPGAGSPGDK-----GETGAPAVGPGSARGAPEGREGSGA 792
QY 194 HSLPEKKNLPDPTSNAT-----SRGABGP.PPPSPMPAVAGAGLALLL 239
D 793 ---PGPAGFAPPGGDDGPGAKGKAGDNGAKGDGAGGAPGFTGAPGPPGAPN----- 843
QY 240 LGVAAGAGACWERRRRRAKPSRHPG-----PSFGRGSLGLGGGGGMPREAPGE 292
D 844 TGAKKARGA-----AGPGATGFPGAAGRPPGSPGNNPPTGPGGKGGKGNRGE 897
QY 293 LGIALRGG--GAADPPFCPHYKVSQDYGHPVYIVODGPPQSP 333
D 898 TGPAGRPGLGAAGP-----GPXGKGGQPGGDPNPGSGTP 934

```

```

RESULT 42
Q90YC5
ID Q90YC5 PRELIMINARY; PRT; 219 AA.
AC Q90YC5;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 20, Last annotation update)
DE Ephrin-A3.
GN EPHRIN-A3.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21412237; PubMed=11520665;
RA Hirate Y., Mieda M., Harada T., Yamasu K., Okamoto H.;
RT "Identification of ephrin-A3 and novel genes specific to the midbrain-
RT MB in embryonic zebrafish by ordered differential display.";

```



```

RESULT 45
O97405 ID O97405 PRELIMINARY; PRT; 1378 AA.
AC O97405;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Collagen pro alpha-chain precursor.
GN HCOL 1 ALPHA.
OS Haliotis discus (Abalone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Archaeogastropoda;
OC Haliotidae; Haliotis.
CC NCBI_TaxID=36094;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE;
RX MEDLINE=99234051; PubMed=10215888;
RA Yoneda C., Hirayama Y., Nakaya M., Matubara Y., Irie S., Hatce K.,
RA Matabe S.;
RT "The occurrence of two types of collagen proalpha-chain in the abalone
RT Haliotis discus muscle.";
RL Eur. J. Biochem. 261:714-721(1999).
DR EMBL; AB017600; BAA75668.1; -
DR InterPro; IPR000885; Fib collagen_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 17.
DR ProDom; PD000007; Collagen; 1.
DR ProDom; PD002078; Fib collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
DR SMART; SM00214; VWC; 1.
DR Signal.
KM Signal.
SQ SEQUENCE 1378 AA; 132583 MW; 6EEB3A65B52E634 CRC64;
FT SIGNAL. 1 18 POTENTIAL.
Query Match 7.6%; Score 140; DB 5; Length 1378;
Best Local Similarity 25.4%; Pred. No. 0.024;
Matches 96; Conservative 21; Mismatches 125; Indels 136; Gaps 20;
QY 2 GPHSGPGGVN-VGALLLGLVGLSLRPYNN-SANKRFQAEQGVLYLPQIGDR 57
DB 646 GPP--GPPAGAGDAGQGLGMPG---ERPGIRNGPQGNRLTGERG---QDGE- 692
QY 58 LDLCPARP-----PGHSSPNYEFYLYVGAGQRCEAPAPNILLTCDR 106
DB 693 -----PGRPEBAGAPGPGPBGSG-----LVSAKDRGHAAPR----- 726
QY 107 PDLRLFTKFOEYSPNLWGHFRSHHDYIATSDGTR--EGLESIQGVCLTRGKVL 164
DB 727 -----GEPGP-----PGSSGQRGPAGAQGPQTGSGPTGEM 759
QY 165 LRVGSPRGAAPVPRKPYSEVEMERDRGAASLE-----PGKENTPGDP--T 208
DB 760 GQTSGDGDCAKCTGATGAGYFGEAGPTICAPNENGRBKSRSGGIPGNSCTPBDGRA 819
QY 209 SNATSRGAEGPLPPSPMPAVAGAGLALLLLGVAGAGAMCWRRAKPSSESHPG-- 265
DB 820 GPPGSPGAGP-----PGSPGAT-----GLSGDGE--RGETGPPGRSGEPGAPG 862
QY 266 -----PGSGRGGSLGIGGGG---MGPR-----EAERBELGILRAGADPFCPHYE 312
DB 863 MPGLDGAQGERGSPRLNPSPPPGVGFQGERGANGFFGSGEA---GAAGPPGSAQGP 918
QY 313 KVSQDYGHVYIVDGP 330
DB 919 GLRGDNGNDAPGQAGP 936
RESULT 46
O9Y1B4 ID O9Y1B4 PRELIMINARY; PRT; 1450 AA.

```

```

AC O9Y1B4;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Alpha 1 type I collagen.
OC Cynops pyrrhogaster (Japanese common newt).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae; Cynops.
CC NCBI_TaxID=8330;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=REGENERATE FORELIMBS;
RX MEDLINE=99407244; PubMed=10474166;
RA Asahina K., Obara M., Yoshizato K.;
RT "Expression of genes of type I and type II collagen in the formation
RT and development of the blastema of regenerating newt limb.";
RL Dev. Dyn. 216:59-71(1999).
DR EMBL; AB015438; BAA36973.1; -
DR InterPro; IPR000885; Fib collagen_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 18.
DR ProDom; PD000007; Collagen; 1.
DR ProDom; PD002078; Fib collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01208; VWF; UNKNOWN_1.
KM Collagen.
SQ SEQUENCE 1450 AA; 137564 MW; ABF8A74841B97B7C CRC64;
FT SIGNAL.
Query Match 7.5%; Score 139.5; DB 13; Length 1450;
Best Local Similarity 25.7%; Pred. No. 0.028;
Matches 103; Conservative 24; Mismatches 137; Indels 137; Gaps 24;
QY 2 GPHSGPGGVN-VGALLLGLVGLSLRPYNN-SANKRFQAEQGVLYLPQIGDR 59
DB 552 GPP--GPPAGAGQCS---GVMGFPQPKGAGP-----GKSGER-- 584
QY 60 LCPARPFPSPHSSPNYEFYLYVGAGQRCEAPAPNILLTCDRDLRLFTKFOE 119
DB 585 ---GVAGPPGATGAKGKG-----EAGNQ-----PPPS--GPGSEREQGPASPGQG 630
QY 120 Y--SPNLWGHFRSHHDYIATSDGTREGLESIQG--GVCLTRGKVLVVGSPRGA 175
DB 631 LPGSPGAPGAEAGKPEQG---APGDAGGPPGSPGRGERGFPERG-----GQGPAGAQ 680
QY 176 VPR---KPVSE-----MPMER-----DRGAHSLERG 199
DB 681 GPRGSPGSPGNDAGKAGAGAGAGGGRGPPGLQGMPPGERSAGMPGAKGDKGA---G 735
QY 200 KENLPDPTSNATSRGAEGLPPSPMPAVAG--AAGLALLLLGVAGAGAMCWRRAK 257
DB 736 TGADGAGAKGDG-ARGLGPICFPSPGAPDKGCGPS---GPAQPTGA-----RGS 784
QY 258 PESRHPG-----PGSFRGGSLG---LGGCGMGREABEGELITAL 297
DB 785 PGRGEPGAPGAGICGPPGADGQAGKESGDAPKADAPGAPGAGTGA--PGPANGV 843
QY 298 RCG-----GADPPFCPHYEKVSGDYGHVYIVDGPDPSP 333
DB 844 APGPCTRGAAAGPAGATGPPGAAAGRLGPPGSPGNAGPFGPP 884
RESULT 47
O9UMG6 ID O9UMG6 PRELIMINARY; PRT; 1690 AA.
AC O9UMG6;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Collagen type IV a6 chain.
GN COL4A6.

```

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OK NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-1644 FROM N.A.
 RX MEDLINE=9629642; PubMed=8661006;
 RA Zhang X., Zhou J., Reedeers S.T., Tryggvason K.;
 RT "Structure of the human type IV collagen COL4A6 gene, which is mutated
 in Alport syndrome-associated leiomyomatosis.";
 RL Genomics 33:473-479(1996).
 RN- [2]
 RP SEQUENCE FROM N.A.
 RA Zhang X.;
 RL Submitted (JAN-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL, U47004; AAB19039.1; -. JOINED.
 DR EMBL, U46960; AAB19039.1; JOINED.
 DR EMBL, U46961; AAB19039.1; JOINED.
 DR EMBL, U46962; AAB19039.1; JOINED.
 DR EMBL, U46963; AAB19039.1; JOINED.
 DR EMBL, U46964; AAB19039.1; JOINED.
 DR EMBL, U46965; AAB19039.1; JOINED.
 DR EMBL, U46966; AAB19039.1; JOINED.
 DR EMBL, U46967; AAB19039.1; JOINED.
 DR EMBL, U46968; AAB19039.1; JOINED.
 DR EMBL, U46969; AAB19039.1; JOINED.
 DR EMBL, U46970; AAB19039.1; JOINED.
 DR EMBL, U46971; AAB19039.1; JOINED.
 DR EMBL, U46972; AAB19039.1; JOINED.
 DR EMBL, U46973; AAB19039.1; JOINED.
 DR EMBL, U46974; AAB19039.1; JOINED.
 DR EMBL, U46975; AAB19039.1; JOINED.
 DR EMBL, U46976; AAB19039.1; JOINED.
 DR EMBL, U46977; AAB19039.1; JOINED.
 DR EMBL, U46978; AAB19039.1; JOINED.
 DR EMBL, U46979; AAB19039.1; JOINED.
 DR EMBL, U46980; AAB19039.1; JOINED.
 DR EMBL, U46981; AAB19039.1; JOINED.
 DR EMBL, U46982; AAB19039.1; JOINED.
 DR EMBL, U46983; AAB19039.1; JOINED.
 DR EMBL, U46984; AAB19039.1; JOINED.
 DR EMBL, U46985; AAB19039.1; JOINED.
 DR EMBL, U46986; AAB19039.1; JOINED.
 DR EMBL, U46987; AAB19039.1; JOINED.
 DR EMBL, U46988; AAB19039.1; JOINED.
 DR EMBL, U46989; AAB19039.1; JOINED.
 DR EMBL, U46990; AAB19039.1; JOINED.
 DR EMBL, U46991; AAB19039.1; JOINED.
 DR EMBL, U46992; AAB19039.1; JOINED.
 DR EMBL, U46993; AAB19039.1; JOINED.
 DR EMBL, U46994; AAB19039.1; JOINED.
 DR EMBL, U46995; AAB19039.1; JOINED.
 DR EMBL, U46996; AAB19039.1; JOINED.
 DR EMBL, U46997; AAB19039.1; JOINED.
 DR EMBL, U46998; AAB19039.1; JOINED.
 DR EMBL, U46999; AAB19039.1; JOINED.
 DR EMBL, U47000; AAB19039.1; JOINED.
 DR EMBL, U47001; AAB19039.1; JOINED.
 DR EMBL, U47002; AAB19039.1; JOINED.
 DR EMBL, U47003; AAB19039.1; JOINED.
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR001442; ProcollagenC4.
 DR Pfam: PF01413; C4; 2.
 DR Pfam: PF01391; Collagen; 23.
 DR ProDom: PD000007; Collagen; 4.
 DR ProDom: PD003923; ProcollagenC4; 1.
 DR SMART; SM00111; C4; 2.
 KW Collagen.
 SQ SEQUENCE 1690 AA; 163696 MW; 4698A2CC2D3E859 CRC64;

QY 147 GLESLQ--GYCLTRKMKVLL-----RVGSPRGA-----VRKPYSEM 184
 Db 905 GFGPIPLPGISGTRKIKGI.PSGTGKMGPSGRAGTPEKGDNRNPGVGI.PSPRRPMSNL 964
 QY 185 PMERDRGAASL-----EPKENTLPDDPTSNATSRGAE-PLP----- 222
 Db 965 WLKDKSSQGSAGSNRPGPRGDKGKNGRPGPLGAPPLPILIKVSGKPGPPGMGI 1024
 QY 223 -----PMPVAVAGAG-----GL--ALLIGVAGAGCMRRRRAXPSE 260
 Db 1025 RGLPGLKSGSGITGFGPMPFSGSGIGRSGPLPGASGLPGLKGNQGV- -EISGSPGP 1082
 QY 261 SRHPGPGSF--GRGSGTGLGGGGMGPREAPRGLGIALRGGAADPPGCHYKVGSDYG 319
 Db 1083 KGQPSGSGFKGTGRDGLGNIGPFGKKGEDKVGVS---GVLGLPAGPFGVAGMRG 1138
 QY 320 HP 321
 Db 1139 EP 1140
 RESULT 48
 ID QY4L4 PRELIMINARY; PRT; 1691 AA.
 AC QY4L4;
 DT 01-NOV-1999 (T-EMBLrel. 12, Created)
 DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE Collagen type IV a6 chain.
 GN COL4A6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OK NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-1645 FROM N.A.
 RX MEDLINE=9629642; PubMed=8661006;
 RA Zhang X., Zhou J., Reedeers S.T., Tryggvason K.;
 RT "Structure of the human type IV collagen COL4A6 gene, which is mutated
 in Alport syndrome-associated leiomyomatosis.";
 RL Genomics 33:473-479(1996).
 RN- [2]
 RP SEQUENCE FROM N.A.
 RA Zhang X.;
 RL Submitted (JAN-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL, U47004; AAB19038.1; -. JOINED.
 DR EMBL, U46958; AAB19038.1; JOINED.
 DR EMBL, U46959; AAB19038.1; JOINED.
 DR EMBL, U46960; AAB19038.1; JOINED.
 DR EMBL, U46961; AAB19038.1; JOINED.
 DR EMBL, U46962; AAB19038.1; JOINED.
 DR EMBL, U46963; AAB19038.1; JOINED.
 DR EMBL, U46964; AAB19038.1; JOINED.
 DR EMBL, U46965; AAB19038.1; JOINED.
 DR EMBL, U46966; AAB19038.1; JOINED.
 DR EMBL, U46967; AAB19038.1; JOINED.
 DR EMBL, U46968; AAB19038.1; JOINED.
 DR EMBL, U46969; AAB19038.1; JOINED.
 DR EMBL, U46970; AAB19038.1; JOINED.
 DR EMBL, U46971; AAB19038.1; JOINED.
 DR EMBL, U46972; AAB19038.1; JOINED.
 DR EMBL, U46973; AAB19038.1; JOINED.
 DR EMBL, U46974; AAB19038.1; JOINED.
 DR EMBL, U46975; AAB19038.1; JOINED.
 DR EMBL, U46976; AAB19038.1; JOINED.
 DR EMBL, U46977; AAB19038.1; JOINED.
 DR EMBL, U46978; AAB19038.1; JOINED.
 DR EMBL, U46979; AAB19038.1; JOINED.
 DR EMBL, U46980; AAB19038.1; JOINED.
 DR EMBL, U46981; AAB19038.1; JOINED.
 DR EMBL, U46982; AAB19038.1; JOINED.
 DR EMBL, U46983; AAB19038.1; JOINED.
 DR EMBL, U46984; AAB19038.1; JOINED.
 DR EMBL, U46985; AAB19038.1; JOINED.

Query Match 7.5%; Score 139.5; DB 4; Length 1690;
 Best Local Similarity 25.2%; Pred. No. 0.034;
 Matches 61; Conservative 25; Mismatches 83; Indels 73; Gaps 11;

```

DR EMBL; U46986; AAB19038.1; JOINED.
DR EMBL; U46987; AAB19038.1; JOINED.
DR EMBL; U46988; AAB19038.1; JOINED.
DR EMBL; U46989; AAB19038.1; JOINED.
DR EMBL; U46990; AAB19038.1; JOINED.
DR EMBL; U46991; AAB19038.1; JOINED.
DR EMBL; U46992; AAB19038.1; JOINED.
DR EMBL; U46993; AAB19038.1; JOINED.
DR EMBL; U46994; AAB19038.1; JOINED.
DR EMBL; U46995; AAB19038.1; JOINED.
DR EMBL; U46996; AAB19038.1; JOINED.
DR EMBL; U46997; AAB19038.1; JOINED.
DR EMBL; U46998; AAB19038.1; JOINED.
DR EMBL; U46999; AAB19038.1; JOINED.
DR EMBL; U47000; AAB19038.1; JOINED.
DR EMBL; U47001; AAB19038.1; JOINED.
DR EMBL; U47002; AAB19038.1; JOINED.
DR EMBL; U47003; AAB19038.1; JOINED.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR001442; ProcollagenC4.
DR Pfam; PF01413; C4; 2.
DR Pfam; PF01391; Collagen; 23.
DR ProDom; PD000007; Collagen; 4.
DR ProDom; PD003923; ProcollagenC4; 1.
DR SMART; SM00111; C4; 2.
KW COLLAGEN.
SQ SEQUENCE. 1691 AA; 163873 MW; 769AA53D1C7CA87 CRC64;

Query Match 7.5%; Score 139.5; DB 4; Length 1691;
Best Local Similarity 25.2%; Pred. No. 0.034;
Matches 61; Conservative 25; Mismatches 83; Indels 73; Gaps 11;

QY 147 GLESLQG--GYCLTRGKMTVL---RVGSGPRGA-----VPRKPYSEM 184
DB 906 GPPGIPGIPGISTGKLGKIPGISTGKMGSPGAGTSGEKDGRNGPVGILPSPRRPSNL 965
QY 185 PMERDRGAHSP-----EPKENVLPDPTSNATSRGAG-PLPP----- 222
DB 966 WLKGDKSGSGSAGSNVGPFGPDGKGEAGRPDPPLGAPLPGITIKVSGKPGPFGMG 1205
QY 223 -----PSMPAVAGAAG---GL--ALLILGVAGAGAMCMRRRAKPS 260
DB 1026 RGLPGLKSSSGITGFPFGPBGSGSGGIRGSDGLPGASGLPLKIDNQTV--EISGSPGP 1083
QY 261 SRHPGPGSF--GRGSLGLGSGGGMFPRAEPDELGIALRGGAADPPFCPHYKVSQDYG 319
DB 1084 KGQPESEGFRTKGRDGLIGNIGFPGKKGEDGKTVS---GDVGLPGAPGFPVAGMRG 1139
QY 320 HP 321
DB 1140 EP 1141

RESULT 49
Q8TEJ5 PRELIMINARY; PRT; 705 AA.
AC Q8TEJ5;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE F1400201 protein (Fragment).
GN F1400201.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
RX NCBI_Taxid=9606;
RY [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RA Tikuva H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL; AK074129; BAB84955.1; -.
FT NON TER 1
SQ SEQUENCE 705 AA; 67430 MW; 07DB85A65A948ED3 CRC64;

Query Match 7.5%; Score 139; DB 4; Length 705;
Best Local Similarity 31.4%; Pred. No. 0.014;
Matches 66; Conservative 9; Mismatches 81; Indels 54; Gaps 11;

QY 159 RGKVTLLRVGQ-----SPRGAVPRKPVSEMERDRGAHSLBPKENLPDPTSNATS 213
DB 204 RGLKKGNGVQGGQGLPGAPGQGA-PGP---GLP-----GPAGLGKXGLDGLPAPDPDKES 256
QY 214 -----RGAGPLPPEPMAYV--GAAGLALLILGVAGAGAMCMRRRAKPS 258
DB 257 GPPGVGPPRGDEGAVGPKPPGVDDVGVVGAAG-----LPGQGSAGKAGBPGTRGPPGL 311
QY 259 -----SESRHGP--GSPFGGSLGLGSGGGMFPRAEPDELGIALRGGAADPPFCPHY 311
DB 312 IEPITGKMGPLGPKQDGRPAVPLDLDGRGEPGDEGFBEGPBGGLGSG---PPLGLGS 367
QY 312 EKVSQDYGHPIVIVDDP-----PQSPNT 336
DB 368 AGLPGRGPP-----GPKGEAGPGGPGV 391

RESULT 50
Q9D2V4 PRELIMINARY; PRT; 744 AA.
AC Q9D2V4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Procollagen, type VIII, alpha 1.
GN COL8A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_Taxid=10090;
RY [1]
RP SEQUENCE FROM N.A.
RC STRINE=57BL/6J; TISSUE=KIDNEY;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukumishi Y., Komoto H., Adachi U., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Mateno Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.T., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
RA Lyons P., Marchionni L., Washima J., Wazarelli U., Wombers P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Stoch K.F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK018742; BAB1385.1; -.
DR MGD; MGI:88463; Col8a1.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 7.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
SQ SEQUENCE 744 AA; 73581 MW; C659BDCBCEDEB9C CRC64;

```

Query Match 7.5%; Score 139; DB 11; Length 744;
Best Local Similarity 23.2%; Pred.No. 0.014;
Matches 83; Conservative 28; Mismatches 104; Indels 142; Gaps 18;

```
QY 67 PPGPHSSPNYEFYKYLIVG-----GAQGR-CEAPPAPNLLTCDRDPDLRFT 114
      |||||
Db 199 PPGPHGLPG-----IGKPGGPGLPQPGAKGERGPKGPQP----- 234
      |||||

QY 115 IKQEQYSPNLMGHEFRSHDYIIATSDGTR---EGLESLOG--GVCLTRGMKVLRYG 168
      |||||
Db 235 -----PGLQGPX-----GKKGMPGLPGLKPPGPMHGPFGVPVGLPGVG 273
      |||||

QY 169 Q-----SPRGAVPRKPVSEMPMERDRGAHSLERKENTLPDP--TSN 210
      |||||
Db 274 KPGVTGFPQGPGLGKPGEPGERGPGQLGVPGVQGPFGMAGVGRGQDGTFGQGFPGG 333
      |||||

QY 211 ATSRGAEGLPPSPMPAVA-----GAAGLALL-----IGVAGAGAMCMR 252
      |||||
Db 334 KGEQGLPGLPGPGLPGVGKPGFPKGDGIGVPGVGLPGRGEKPGIGAPGMGQ----- 388
      |||||

QY 253 RRRAKPSSESRHPG-PGSPRGGSIGL-----GGGGGWMGPR-----EAPGELGI----- 295
      |||||
Db 389 -----PGEPLPGIIPGMPGPAIGFPGKGEQGVVFGPPGPKGEPGLQGFPGKPGFL 444
      |||||

QY 296 -----ALRG-----GAAADPPFCPHYEKVSGDYGHVPYIVQDGPPOSPPNI 336
      |||||
Db 445 GEVGPMPMRGLPGPIGPKGEGHKGLPGLPGVGLLGPKEP-GIPGQGLQGPPI 500
      |||||
```

Search completed: February 11, 2003, 12:05:05
Job time : 36.2327 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: February 11, 2003, 11:47:24 ; Search time 14.5409 Seconds
(without alignments)
969.814 Million cell updates/sec

Title: US-10-021-121-4

Perfect score: 1850
Sequence: 1 MGPPHSGPGVAVGALLILG.....PVIIVDGPSPGSPNITYKV 340

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476326 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database : SWISSPROT_40:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1850	100.0	340	1 EFB3_HUMAN	Q15768 homo sapien
2	1780	96.2	340	1 EFB3_MOUSE	O35393 mus musculu
3	637.5	34.5	336	1 EFB2_MOUSE	P52800 mus musculu
4	632	34.2	346	1 EFB1_HUMAN	P98172 homo sapien
5	629.5	34.0	333	1 EFB2_HUMAN	P52799 homo sapien
6	628	33.9	334	1 EFB1_CHICK	O73612 gallus galli
7	626.5	33.9	332	1 EFB2_BRARE	O73874 brachydacti
8	613.5	33.2	345	1 EFB1_MOUSE	P52795 mus musculu
9	608.5	32.9	345	1 EFB1_RAT	P52796 rattus norv
10	600	32.4	327	1 EFB1_XENLA	O13097 xenopus lae
11	185	10.0	195	1 EFB2_BRARE	P79727 brachydacti
12	179	9.7	238	1 EFB3_HUMAN	P52797 homo sapien
13	176	9.5	209	1 EFB2_MOUSE	P52801 mus musculu
14	175.5	9.5	213	1 EFB2_HUMAN	O43921 homo sapien
15	172	9.3	200	1 EFB2_CHICK	P52802 gallus galli
16	170.5	9.2	228	1 EFB5_CHICK	P52804 gallus galli
17	169.5	9.2	201	1 EFB4_HUMAN	P52798 homo sapien
18	169.5	9.2	228	1 EFB5_BRARE	P79728 brachydacti
19	167.5	9.1	216	1 EFB1_XENLA	P52794 xenopus lae
20	167.5	9.1	228	1 EFB5_HUMAN	P52803 homo sapien
21	167.5	9.1	228	1 EFB5_MOUSE	O08543 mus musculu
22	167.5	9.1	228	1 EFB5_RAT	P97605 rattus norv
23	166	9.0	205	1 EFB1_HUMAN	P20827 homo sapien
24	162.5	8.8	205	1 EFB1_RAT	P97553 rattus norv
25	161	8.7	205	1 EFB1_MOUSE	P52793 mus musculu
26	160.5	8.7	206	1 EFB4_MOUSE	O08542 mus musculu
27	159	8.6	680	1 CALA_MOUSE	O08306 mus musculu
28	154.5	8.4	1049	1 CAL3_BOVIN	P04258 bos taurus
29	151.5	8.2	301	1 CC02_CAEEL	P16565 caenorhabdi
30	148.5	8.0	1670	1 CAL4_HUMAN	O01955 homo sapien
31	146.5	7.9	1527	1 CALH_MOUSE	P30765 mus musculu
32	146	7.9	1027	1 CALF_RIFPA	P30754 riftia pach
33	145	7.8	674	1 CALA_BOVIN	P23206 bos taurus

34	144.5	7.8	1516	1 CALH_HUMAN	P39060 homo sapien
35	144	7.8	675	1 CA39_CHICK	P32017 gallus galli
36	144	7.8	921	1 CAL9_HUMAN	P20849 homo sapien
37	143	7.7	674	1 CALA_CHICK	P08125 gallus galli
38	142.5	7.7	1496	1 CA25_HUMAN	P05997 homo sapien
39	142	7.7	744	1 CAL8_HUMAN	P27658 homo sapien
40	142	7.7	1029	1 CAL6_MOUSE	O02788 mus musculu
41	142	7.7	1763	1 CA24_ASCSU	P27393 ascaris suu
42	141.5	7.6	1466	1 CAL3_HUMAN	P02461 homo sapien
43	141	7.6	744	1 CAL8_RABIT	P14282 oryctolagus
44	140.5	7.6	743	1 CAL8_MOUSE	O00780 mus musculu
45	140	7.6	680	1 CALA_HUMAN	O03692 homo sapien
46	139.5	7.5	200	1 CAL2_HAEKO	P16252 haemonchus
47	139	7.5	1804	1 CALB_MOUSE	O61245 mus musculu
48	138.5	7.5	1459	1 CAL2_MOUSE	P28481 mus musculu
49	138.5	7.5	1736	1 CA2B_HUMAN	P13942 homo sapien
50	138	7.5	1758	1 CA24_CAEEL	P17140 caenorhabdi

ALIGNMENTS

```

RESULT 1
ID EFB3_HUMAN STANDARD; PRT; 340 AA.
AC Q15768; O06880; Q92875;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ephrin-B3 precursor (EPH-related receptor tyrosine kinase ligand 8)
DE (EPRK-8) (EPH-related receptor transmembrane ligand EPRK-12).
GN EFB3 OR EPRK8 OR EPRK8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RX NCBI_Taxid=9606;
RN [1]
RP SOURCE FROM N.A.
RA Cerretti D.P.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RP [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96404527; PubMed=9126477;
RX MEDLINE=97271551; PubMed=9126477;
RA Tang X.X., Pleasure D.E., Ikegaki N.;
RT "CDNA cloning, chromosomal localization, and expression pattern of
RT EPRK8, a new member of the EPRK gene family encoding ligands of EPH-
RT related protein-tyrosine kinase receptors.";
RL Genomics 41:17-24(1997).
RN [3]
RP SOURCE FROM N.A.
RC TISSUE=Brain cortex;
RX MEDLINE=96404527; PubMed=8808709;
RA Gale N.W., Fleniken A., Compton D.C., Jenkins N.A., Copeland N.G.,
RA Giallardi D.J., Davis S., Wilkinson D.G., Yancopoulos G.D.;
RT "EPRK-13, a novel transmembrane ligand for the Eph family of receptor
RT tyrosine kinases, expressed in embryonic floor plate, roof plate and
RT hindbrain segments.";
RL Oncogene 13:1343-1352(1996).
CC -1- FUNCTION: MAY PLAY A PIVOTAL ROLE IN FOREBRAIN FUNCTION. BINDS TO,
CC AND INDUCE THE COLLAPSE OF, COMMISSURAL AXONS/GROWTH CONES IN
CC VITRO. MAY PLAY A ROLE IN CONSTRAINING THE ORIENTATION OF
CC LONGITUDINALLY PROJECTING AXONS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN; EXPRESSED IN
CC EMBRYONIC FLOOR PLATE, ROOF PLATE AND HINDRAIN SEGMENTS.
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```


DB 241 GVAGAGCAGCRRRRAXPESRRHGPSPFRGSGSLGLGGGGGNGPREAREPGLGIALRG 300
 QY 301 GAADPPFCPPHYEKVSGDYGHPPVTVODGPPSPFNITYKV 340
 DB 301 GTADPPFCPPHYEKVSGDYGHPPVTVODGPPSPFNITYKV 340

RESULT 3
 EFB2_MOUSE STANDARD; PRT; 336 AA.
 ID EFB2_MOUSE
 AC P52800;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ephrin-B2 precursor (EPH-related receptor tyrosine kinase ligand 5)
 DE (LEK-5) (HTK ligand) (HTK-L) (ELF-2)
 GN EPNB2 OR EPLG5 OR LEK5 OR HTKL OR ELF2 OR EPL5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96145238; PubMed=8559144;
 RA Cerretti D.P., Vanden Bos T., Nelson N., Kozlosky C.J., Reddy P.,
 RA Maraskovsky E., Park L.S., Lyman S.D., Copeland N.G.,
 RA Gilbert D.J., Jenkins N.A., Fletcher R.A.;
 RA "Isolation of LEK-5: a ligand of the eph-related receptor tyrosine
 RT kinases.";
 RT Mol. Immunol. 32:1197-1205(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CB57BL/6J X SJL/J;
 RX MEDLINE=95199254; PubMed=7534404;
 RA Bennett B.D., Zeigler F.C., Gu Q., Fendly B., Goddard A.D.,
 RA Gillett N., Mathews W.;
 RT "Molecular cloning of a ligand for the EPH-related receptor protein-
 RT tyrosine kinase Hck.";
 RT Proc. Natl. Acad. Sci. U.S.A. 92:1866-1870(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR; TISSUE=Brain;
 RX MEDLINE=95379837; PubMed=7651410;
 RA Bergemann A.D., Cheng H.-J., Brambilla R., Klein R., Flanagan J.G.;
 RT "ELF-2, a new member of the Eph ligand family, is segmentally
 RT expressed in mouse embryos in the region of the hindbrain and newly
 RT forming somites.";
 RT Mol. Cell. Biol. 15:4921-4929(1995).
 RN [4]
 RP FUNCTION.
 RX MEDLINE=20171264; PubMed=10704386;
 RA Imondi R., Wideman C., Kaprielian Z.;
 RT "Complementary expression of a transmembrane ephrins and their receptors
 RT in the mouse spinal cord: a possible role in constraining the
 RT orientation of longitudinally projecting axons.";
 RT Development 127:1397-1410(2000).
 CC -1- FUNCTION: MAY PLAY A ROLE IN CONSTRAINING THE ORIENTATION OF
 CC LONGITUDINALLY PROJECTING AXONS.
 CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASE EPHB4.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED ON LATERAL FLOOR PLATE CELLS,
 CC SPECIFICALLY ON COMMISSURAL AXON SEGMENTS THAT HAVE PASSED THROUGH
 CC THE FLOOR PLATE. EXPRESSED IN CELLS OF THE RETINAL GANGLION CELL
 CC LAYER DURING RETINAL AXON GUIDANCE TO THE OPTIC DISC.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE FLOOR PLATE THROUGHOUT THE
 CC PERIOD OF COMMISSURAL AXON PATHFINDING.
 CC -1- PTM: INDUCIBLE PHOSPHORYLATION OF TYROSINE RESIDUES IN THE
 CC CYTOPLASMIC DOMAIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U16819; AAA99708.1; -
 DR EMBL; L38847; AAC42052.1; -
 DR EMBL; U30244; AAA82934.1; -
 DR MGD; MG1:105097; Efnb2.
 DR InterPro; IPR001799; Ephrin.
 DR Pfam; PF00812; Ephrin; 1.
 DR PRINTS; PR01347; EPHRIN.
 DR ProDom; PD002533; Ephrin; 1.
 DR PROSITE; PS01299; EPHRIN; 1.
 KW Developmental protein; Neurogenesis; Transmembrane; Glycoprotein;
 KW Signal; Phosphorylation.
 FT SIGNAL 1 28
 FT CHAIN 29 336
 FT DOMAIN 29 232
 FT TRANSMEM 233 253
 FT DOMAIN 254 336
 FT DOMAIN 334 336
 FT CARBOHYD 339 339
 FT CARBOHYD 142 142
 FT CONFLICT 3 4
 FT CONFLICT 177 177
 SQ SEQUENCE 336 AA; 37202 MW; D0889496E39554 CRC64;
 Query Match 34.5%; Score 637.5; DR 1; Length 336;
 Best Local Similarity 42.1%; Pred. No. 1,1e-34;
 Matches 139; Conservative 49; Mismatches 129; Indels 13; Gaps 5;
 QY 14 GATLLGLVGLVSGLSLEPYNNANRRFOAGGYVLPQIGDRDLDCPRAPGPBHS 73
 DB 17 GLAMVLCRTAISRSIVLEPIYNNSSNEKFLPGQGLVLYPGIGKLDICPKV--DSKTV 73
 QY 74 PNYEFLYLYVGAGQRCGAPAPANNLLTCDPDLRTFTFOEYSPMLMHEPRSH 133
 DB 74 GQYEVYVYVWDKQADRCITKENTPLNLCARPDDQVKTFIQEESPMLWGIPOKX 133
 QY 134 DYYIIATSDGREGLESLGAGVCTIRGMKVLIRVQ--SPRGAVPKPPSEMPMER-DR 130
 DB 134 DYYIISTNSGSLBDLNDQBGVCTIRAMKIMLVQDASSAGABRNHGPFRPELEAGTN 193
 QY 191 GAHSLEPGKENTPGDPTSNATSRGAEGLPPSPMAVAGAGLALLIGVAGAGAMC 250
 DB 194 GRSSITPPFYVFNPGSSITDNSAGHSNNLIGSEVALFAGIASGCIFIVYIITLVVLL 253
 QY 251 WRRRRAPSESRRHGPSPFRGSGSLGLGGGGGNGPREAREPGLGIALRGGAADPPCPH 310
 DB 254 KYRRRRHKSPEQHTTTLSTLTATPKRGKNN---NGSBSVDVITPLR---TADSVCPH 306
 QY 311 YEKVSGDYGHPPVTVODGPPSPFNITYKV 340
 DB 307 YEKVSGDYGHPPVTVODGPPSPFNITYKV 336

RESULT 4
 EFB1_HUMAN STANDARD; PRT; 346 AA.
 ID EFB1_HUMAN
 AC P98172;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ephrin-B1 precursor (EPH-related receptor tyrosine kinase ligand 2)
 DE (LEK-2) (ELK ligand) (ELK-L).
 GN EPNB1 OR EPLG2 OR LEK2 OR ELF-3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RC TISSUE=Placenta; PubMed=8070404;
RX Beckmann M.P., Czeretel D.P., Baum P., Vandenbos T., James L.,
RA Farrah T., Kozlosky C., Hollingsworth T., Shilling H., Maraskovsky E.,
RA Fletcher F.A., Lhotek V., Pawson T., Lyman S.D.;
RT "Molecular characterization of a family of ligands for eph-related
RT tyrosine kinase receptors."
RL EMBO J. 13:3757-3762(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Davis S., Gale N.W., Aldrich T.H., Maisompierre P.C., Lhotek V.,
RA Pawson T., Goldfarb W., Yancopoulos G.D.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Fletcher F.A., Huebner K., Shaffer L.G., Monaco A., Mueller U.,
RA Kozlosky C., Druck T., Simoneaux D.K., Fairweather N., Chelly J.,
RA Czeretel D.P., Belmont J.W., Beckmann M.P., Lyman S.D.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Howden P.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: BINDS TO, AND INDUCE THE COLLAPSE OF, COMMISSURAL
CC AXONS/GROWTH CONES IN VITRO. MAY PLAY A ROLE IN CONSTRAINING THE
CC ORIENTATION OF LONGITUDINALLY PROJECTING AXONS (BY SIMILARITY).
CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHB1 AND EPHA1.
CC BINDS GRIP1 AND GRIP2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: HEART, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE,
CC KIDNEY, PANCREAS.
CC -1- INDUCTION: BY TNF-ALPHA.
CC -1- PMW: INDICIBLE PHOSPHORYLATION OF TYROSINE RESIDUES IN THE
CC CYTOPLASMIC DOMAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL; U09304; AAA53093.1; -
DR EMBL; U37361; AAA52369.1; -
DR EMBL; U09303; AAB41127.1; -
DR EMBL; AL136092; CAB86409.1; -
DR Genew; HGNC:3226; EPHB1.
DR MIM; 300035; -
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin; 1.
DR PRINTS; PRO1347; EPHRIN.
DR PRODOM; PD002533; Ephrin; 1.
DR PROSITE; PS01299; EPHRIN; 1.
KW Developmental protein; Neurogenesis; Transmembrane; Glycoprotein;
KW Signal; Phosphorylation.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 346 EPHRIN-B1.
FT DOMAIN 25 237 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 238 258 POTENTIAL.
FT DOMAIN 259 346 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 344 346 PDZ RECOGNITION MOTIF (POTENTIAL).
FT CARBOHYD 139 139 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 346 AA; 38006 MW; 473DD2F1A5B89DE CRC64;
Query Match 34.2%; Score 632; DB 1; Length 346;
Best Local Similarity 39.5%; Pred. No. 2.5e-34;
Matches 145; Conservative 48; Mismatches 116; Indels 58; Gaps 9;
Cy 8 PGQVAVGALLLGVGVSGV-----SLPEYVNSANKRQAGGVYVQIGSDRLDL 61
Db 4 PGQRWLGKVLAVMVMVWALCRLATPLAKVLEFVSWSSLNPKFLSGKGLVYIPKIGDKDII 63

```

```

QY 62 CPAPRPPGHSSPNYEFYLYLVGAQGRCPAPAPNLLTCDSPDIDRTIFKQEYS 121
Db 64 CPAPRGRP-----YEVYLYLVVREPDAAACSTVDPNVLYCNPEQIRFTIFQES 118
QY 122 PNLWGHFRRSHHDYIYIATSDGTREGLSLQGVCTLRGMKYLRFVQGS PRGAAVRKV 181
Db 119 PNYMGLFEPKHHDIYITSTNSNLEGLNREBGVCRTRTMKIMKVGDDPNVATPEQLTT 178
QY 182 SEMPERDRGAHSLP-PCKENLPDGPITSNATSRAGELPLPPSPAYVAAAGCLA----- 236
Db 179 SRRSRKADVTVCATQAPESRSLDSDSKHRTVNOEKSGP-----GASGSSGDPD 231
QY 237 -----LLLVGAAGGA-----MCFRRRKPSESHPGSGFRGSGSLG 277
Db 232 GFNRSKVALFAVAGACVIFLLIIFLYVLLKRRKRRTQO-----RAALSL 282
QY 278 ---GGGGMGSRERAPGELGIALRGGAADPPCFPHYKXSGDYGHVYIVQDPPQSP 333
Db 283 SLTASPGSGGTAGTEPSPDIIPLR---TTEVNYCPHYEKVSGDYGHVYIVQEMPQSP 339
QY 334 PNIVYKV 340
Db 340 ANIVYKV 346
RESULT 5
EFB2 HUMAN STANDARD; PRT; 333 AA.
ID EFB2 HUMAN
AC P52759;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ephrin-B2 precursor (EPH-related receptor tyrosine kinase ligand 5)
DE (LEKRS-5) (HTK ligand) (HTK-L).
GN EFB2 OR EBI5 OR LERKS OR HTKL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96145238; PubMed=8559144;
RA Czeretel D.P., Vanden Bos T., Nelson N., Kozlosky C.J., Reddy P.,
RA Maraskovsky E., Park U.S., Lyman S.D., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A., Fletcher R.A.;
RT "Isolation of LERK-5; a ligand of the eph-related receptor tyrosine
RT kinases."
RL Mol. Immunol. 32:1197-1205(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95199254; PubMed=7534404;
RA Bennett B.D., Zeigler F.C., Gu Q., Fendly B., Goddard A.D.,
RA Gillett N., Matthews W.;
RT "Molecular cloning of a ligand for the EPH-related receptor protein-
RT tyrosine kinase Hck."
RL Proc. Natl. Acad. Sci. U.S.A. 92:1866-1870(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98192220; PubMed=9533549;
RA Vogt T., Stolz W., Welsh J., Jung B., Kerbel R.S., Kobayashi H.,
RA Landthaler M., McClelland M.;
RT "Overexpression of Lerk-5/Ep15 messenger RNA: a novel marker for
RT increased tumorigenicity and metastatic potential in human malignant
RT melanomas."
RL Clin. Cancer Res. 4:791-797(1998).
CC -1- FUNCTION: MAY PLAY A ROLE IN CONSTRAINING THE ORIENTATION OF
CC LONGITUDINALLY PROJECTING AXONS (BY SIMILARITY).
CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHB4 AND EPHA3.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: LONG AND KIDNEY.

```

CC -1- PTM: INDUCIBLE PHOSPHORYLATION OF TYROSINE RESIDUES IN THE
 CC CYTOPLASMIC DOMAIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U16797; AAC9707.1; -
 CC EMBL: U18724; AAC1752.1; -
 CC EMBL: U18762; AAC03786.1; -
 CC Genew: HGNC:3227; EFNH2.
 CC MIM: 600527; -
 CC InterPro: IPR001799; Ephrin.
 CC Pfam: PF00812; Ephrin.1.
 CC PRINTS: PRO1347; EPHRIN.
 CC ProDom: PD002533; Ephrin.1.
 CC PROSITE: PS01299; EPHRIN.1.
 CC Developmental protein; Neurogenesis; Transmembrane; Glycoprotein;
 CC Signal; Phosphorylation.
 CC FT SIGNAL 1 27 POTENTIAL.
 CC FT CHAIN 28 333 EPHRIN-B2.
 CC FT DOMAIN 28 229 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 230 250 POTENTIAL.
 CC FT DOMAIN 251 333 CYTOPLASMIC (POTENTIAL).
 CC FT PRINTS 331 333 PDZ RECOGNITION MOTIF (POTENTIAL).
 CC FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SQ SEQUENCE 333 AA; 36923 MW; 6D9932A632626A6A CRC64;
 CC
 CC Query Match 34.0%; Score 629.5; DB 1; Length 333;
 CC Best Local Similarity 40.9%; Pred. No. 3,5e-34;
 CC Matches 135; Conservative 52; Mismatches 130; Indels 13; Gaps 5;
 CC
 CC QY 14 GALLUSGVLGVLSGLEPYYNSANKRPOAGGYLYPOIGDRDLDCPRARPPGPHSS 73
 CC DB 14 GVLAVLCRAIAISIVLEPIYNNSSKFLPGQGLVLYPOIGDKLIDCPKV---DSKTV 70
 CC QY 74 PNYEFYKLYLVGAQGRCEAPPAENLLTCDRPLDLRFTTKFOEYSNLMGHFRSHH 133
 CC DB 71 GQYEVYKVMVVDQADRCITKENTPLNCAKPDODIKFTIKFOEFSNLMGLFQNKX 130
 CC QY 134 DYITITSGTRGLESLOGGVCLTRGMVLLRVGQ--SPRGAVPRKVSSEMPMER-DR 190
 CC DB 131 DYIITSTNGSLEGLDNQGGVQOTRAMKITMKVGDASASTRKQPTRRPELEAGTN 190
 CC QY 191 GAHSLPEPKENLPDPTNATSGALEGPLPPSPMPAVGAAGLALLLVGAAGAGAMC 250
 CC DB 191 GNSSTTSPTKRPSSSTGNSAGHGNMILGSEVALFAGISGCIFFVIITLVLL 250
 CC QY 251 WRRRAKPSBSRHPGSGTGRGSLGLGGGKGPREAPBGLALRGGAADPPFCFH 310
 CC DB 251 KXRRRRKRSPOHTTTLISLTATPRRSQNN---NGSEPSDIIIPLR---TADSVFCFH 303
 CC QY 311 YEKVSGDYGHPIYVODGPSPSPNIIYKY 340
 CC DB 304 YEKVSGDYGHPIYVODGPSPNIIYKY 333
 CC
 CC RESULT 6
 CC EFB1 CHICK STANDARD; PRT; 334 AA.
 CC AC 073612;
 CC DT 15-JUL-1999 (Rel. 38, Created)
 CC DT 15-JUL-1999 (Rel. 38, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Ephrin-B1 precursor (CEK5 ligand) (CELS-L).
 CC GN EFNH1.
 CC OS Gallus gallus (Chicken).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 CC NCBI_TaxID=9031;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=9723524; Pubmed=9070326;
 CC RA Holash J.A., Soans C., Chong L.D., Shao H., Dixie V.M.,
 CC Pasquale E.B.;
 CC RT "Reciprocal expression of the Eph receptor Cdk5 and its ligand(s) in
 CC the early retina".
 CC RL Dev. Biol. 182:256-269(1997).
 CC CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASE EPHB2.
 CC CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC CC -1- PTM: INDUCIBLE PHOSPHORYLATION OF TYROSINE RESIDUES IN THE
 CC CYTOPLASMIC DOMAIN (BY SIMILARITY).
 CC CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U72394; AAC07986.1; -
 CC InterPro: IPR001799; Ephrin.
 CC Pfam: PF00812; Ephrin.1.
 CC PRINTS: PRO1347; EPHRIN.
 CC ProDom: PD002533; Ephrin.1.
 CC PROSITE: PS01299; EPHRIN.1.
 CC Developmental protein; Neurogenesis; Transmembrane; Glycoprotein;
 CC Signal; Phosphorylation.
 CC FT SIGNAL 1 25 POTENTIAL.
 CC FT CHAIN 26 334 EPHRIN-B1.
 CC FT DOMAIN 26 231 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 232 252 POTENTIAL.
 CC FT DOMAIN 253 334 CYTOPLASMIC (POTENTIAL).
 CC FT PRINTS 332 334 PDZ RECOGNITION MOTIF (POTENTIAL).
 CC FT CARBOHYD 133 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SQ SEQUENCE 334 AA; 36858 MW; 48AF56B9E56CD5 CRC64;
 CC
 CC Query Match 33.9%; Score 628; DB 1; Length 334;
 CC Best Local Similarity 39.7%; Pred. No. 4,4e-34;
 CC Matches 146; Conservative 50; Mismatches 100; Indels 72; Gaps 13;
 CC
 CC QY 8 PGYVR--VGALLIGVLGVLSGLEPYYNSANKRPOAGGYLYPOIGDRDLDCPR 65
 CC DB 4 PGRGRWMLGVLLALCRLAAPLAKSLEFVSWASAGNPFMSGKGVIVYBEIDDKLIDCPKA 63
 CC QY 66 RPPGSHSPNYEFYKLYLVGAQGRCEAPPAENLLTCDRPLDLRFTTKFOEYSNLM 125
 CC DB 64 EPRKPE---YDYIKLYLVKKDQADACSTMDPNVLTCKRPFQIRFTTKFOEFSNLM 118
 CC QY 126 GHEFRSHDYIATSDGTREGLESLOGGVCLTRGMVLLRVGQSPRGAVPRKVSSEMP 185
 CC DB 119 GLEFRQODFYITSTNGTLDGLENREGVQCTRSKTIWKVQODP--NAVIPQLTTSRP 177
 CC QY 186 MER-----DGAHSL-----EPKKNLPDPTNSA--TSRGAEGEPLPPMPAVAGA 231
 CC DB 178 SKEDANTVKIVTQSPRKHVPTVEBPK--PQSVNONGETQPSDFL--SKVAVFAA 232
 CC QY 232 AGG-----LALLLVGAAGAGAMCRRRAKPSBSR-----PPGSGFRG 272
 CC DB 233 IGAGCVIFILIIIFLVLLIKI-----RKRRKHKTQGAALSLSTLAPFKSGNA 283
 CC QY 273 GSLGSGGSGGKGRBAPBGLGIALRGGAADPPFCFHNYKVSQDYGHPYTYVODGPSP 332
 CC DB 284 GS-----EPSDIIPLR--TTENNYCPHEKVSQDYGHPYTYVODGPSP 326
 CC QY 333 PPNIIYKY 340
 CC |||||

Db 327 PANIYKV 334

RESULT 7

EFB2_BRARE EFB2_BRARE STANDARD; PRT; 312 AA.
 ID EFB2_BRARE
 AC 073874;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ephrin-B2 precursor.
 GN EPHN2 OR EPHN2A.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NCBI_Taxid=7955;
 RX SEQUENCE FROM N.A.
 RX MEDLINE=98438455; PubMed=9765210;
 RA Durbin L., Brennan C., Shiomi K., Cooke J., Barrios A.,
 RA Shanmugasalingam S., Guthrie B., Lindberg R., Holder N.,
 RT "Eph signaling is required for segmentation and differentiation of
 the somites.";
 RT Genes Dev. 12:3096-3109 (1998).
 RL -1 SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASE EPHB4.
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1 PTM: INDUCIBLE PHOSPHORYLATION OF TYROSINE RESIDUES IN THE
 CYTOPLASMIC DOMAIN (BY SIMILARITY).
 CC -1 SIMILARITY: BELONGS TO THE EPHRIN FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 DR EMBL; AJ004863; CA006168.1;
 DR ZFIN; ZDB-GENE-990415-67; efnb2a.
 DR InterPro; IPR001799; Ephrin.
 DR Pfam; PF00812; Ephrin.1.
 DR PRINTS; PR01347; EPHRIN.
 DR PRODOM; PD002533; Ephrin.1.
 DR PROSITE; PS01299; EPHRIN.1.
 KM Developmental protein; Neurogenesis; Transmembrane; Glycoprotein;
 KW Signal; Phosphorylation.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 332 EPHRIN-B2.
 FT DOMAIN 25 225 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 226 246 POTENTIAL.
 FT DOMAIN 247 332 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 330 332 PDZ RECOGNITION MOTIF (POTENTIAL).
 SQ SEQUENCE 332 AA; 36724 MW; 189ED8237C71C8B CRC64;

Query Match 33.9%; Score 626.5; DB 1; Length 332;
 Best Local Similarity 42.2%; Pred. No. 5.4e-34;
 Matches 145; Conservative 54; Mismatches 106; Indels 39; Gaps 12;

QY 14 GALLLGLVGVSGLSLPPVYNSANKRFOAEGGVLYPCIGRLDLLCPARPFGPHSS 73
 11 GVLVIAKVNISRALILDSITVNTTKTFVPGQGLVLYPCIGRLDLLCPARPFGPHSS 67
 QY 74 PNYEYKYLIVGAGRGCEAPAPNLLITCDRPLDLRFITKQVSPNMGHEFFSHR 133
 68 EGVEYKLYMPVLEQLKSCQVTKADTPLLNCVKKPDQVKTLEKQESSPNLMGLEFRGK 127
 QY 134 PYYIATSDGTFEESGSGQGVCLTRGKVLRLVQSPRGAVPRK-PVSEMPERERGA 192
 128 DYIITSTNGMISLNDNEGVCCKSKMIIMKQGNSDPISPKDPTSPRPGPLDG 187
 QY 193 AHS-----LEP-----GKENTPGDPTSNATSRGABGPIPPSPMAVAAGAGLALLILGV 242

Db 188 KDSKSNVTLKPPASPHGDK--GDGNKSSVIGSEVAL-----PACIASVIVIIITML 241

QY 243 AGAGAMCWRRRRAKPPSSRRHPGSGFG-----RGSGLGCGGCGWPREAPEGLIA 296
 242 VFL--ILKRRRRHRKGS-POHATTLISLTATPKRGGS-----GGANNNG---SEPDITIP 291

QY 297 IARGGAADPPFCPEHYEKVSGDGHFVYIVODGPPSPNIIYKV 340
 292 LR---TADSVFCPEHYEKVSGDGHFVYIVODGPPSPNIIYKV 332

RESULT 8

EFB1_MOUSE EFB1_MOUSE STANDARD; PRT; 345 AA.
 ID EFB1_MOUSE
 AC P52755;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Ephrin-B1 precursor (EPH-related receptor tyrosine kinase ligand 2)
 DE (LEKR-2) (ELK ligand) (ELK-1) (STRAL protein) (CEK5 receptor ligand)
 DE (CEK5-L).
 GN EPHB1 OR EPLG3 OR LEKR2 OR STRAL OR EPL2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_Taxid=10090;
 RX SEQUENCE FROM N.A.
 RX STRAIN=129/SV;
 RX MEDLINE=95203867; PubMed=7896266;
 RA Fletcher F.A., Renshaw B., Hollingsworth T., Baum P., Lyman S.D.,
 RA Jenkins N.A., Gilbert D.J., Copeland N.G., Davidson B.L.,
 RT "Genomic organization and chromosomal localization of mouse Eplg2, a
 RT gene encoding a binding protein for the receptor tyrosine kinase
 RT elk.";
 RT Genomics 24:127-132 (1994).

CC
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95377533; PubMed=7649373;
 RA Boullier P., Oulad-Abdelghani M., Vicaire S., Garnier J.M.,
 RA Schaubert B., Dolle P., Chambon P.;
 RT "Efficient cloning of cDNAs of retinoic acid-responsive genes in p19
 RT embryonal carcinoma cells and characterization of a novel mouse gene,
 RT Stral (mouse LEKR-2/Eplg2).";
 RT Dev. Biol. 170:420-433 (1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=95014510; PubMed=7929389;
 RX Shao H., Lou D., Pandey A., Pasquale E.B., Dixit V.M.;
 RT "cDNA cloning and characterization of a ligand for the receptor
 RT protein-tyrosine kinase.";
 RT J. Biol. Chem. 269:26606-26609 (1994).
 RN [4]
 RP FUNCTION.
 RX MEDLINE=20171264; PubMed=10704386;
 RA Imondi R., Wiemann C., Kaprielian Z.;
 RT "Complementary expression of transmembrane ephrins and their receptors
 RT in the mouse spinal cord: a possible role in constraining the
 RT orientation of longitudinally projecting axons.";
 RL Development 127:1397-1410 (2000).
 CC -1 FUNCTION: BINDS TO, AND INDUCE THE COLLAPSE OF, COMMISSURAL
 AXONS/GROWTH CONES IN VITRO. MAY PLAY A ROLE IN CONSTRAINING THE
 ORIENTATION OF LONGITUDINALLY PROJECTING AXONS.
 CC -1 SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHB1 AND EPHA1.
 CC -1 TISSUE SPECIFICITY: EXPRESSED ON LATERAL FLOOR PLATE CELLS.
 CC SPECIFICALLY ON COMMISSURAL AXON SEGMENTS THAT HAVE PASSED THROUGH
 THE FLOOR PLATE. EXPRESSED IN CELLS OF THE RETINAL GANGLION CELL
 LAYER DURING RETINAL AXON GUIDANCE TO THE OPTIC DISC.
 CC -1 DEVELOPMENTAL STAGE: EXPRESSED IN THE FLOOR PLATE THROUGHOUT THE
 PERIOD OF COMMISSURAL AXON PATHFINDING.

CC
 QY 188 KDSKSNVTLKPPASPHGDK--GDGNKSSVIGSEVAL-----PACIASVIVIIITML 241
 243 AGAGAMCWRRRRAKPPSSRRHPGSGFG-----RGSGLGCGGCGWPREAPEGLIA 296
 242 VFL--ILKRRRRHRKGS-POHATTLISLTATPKRGGS-----GGANNNG---SEPDITIP 291
 QY 297 IARGGAADPPFCPEHYEKVSGDGHFVYIVODGPPSPNIIYKV 340
 292 LR---TADSVFCPEHYEKVSGDGHFVYIVODGPPSPNIIYKV 332

CC -1- PTM: INDUCIBLE PHOSPHORYLATION OF TYROSINE RESIDUES IN THE
CC CYTOPLASMIC DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U07602; AAC53247.1; -
CC EMBL: U07598; AAC53247.1; JOINED.
CC EMBL: U07599; AAC53247.1; JOINED.
CC EMBL: U07600; AAC53247.1; JOINED.
CC EMBL: Z48761; CAA8695.1; -
CC EMBL: U12983; AAA53231.1; -
CC MGJ: MGJ102708; Efb1.
CC InterPro: IPR001799; Ephrin.
CC Pfam: PF00812; Ephrin.1.
CC PRINTS: PRO1347; EPHRIN.
CC ProDom: PD002533; Ephrin.1.
CC PROSITE: PS01299; EPHRIN.1.
CC Developmental protein; Neurogenesis; Transmembrane; Glycoprotein;
CC Signal; Phosphorylation.
CC FT CHAIN 1 24 POTENTIAL.
CC FT DOMAIN 25 345 EPHRIN-B1.
CC FT TRANSMEM 237 257 EXTRACELLULAR (POTENTIAL).
CC FT DOMAIN 258 345 POTENTIAL.
CC FT DOMAIN 343 345 CYTOPLASMIC (POTENTIAL).
CC FT CARBOHYD 139 139 PDZ RECOGNITION MOTIF (POTENTIAL).
CC FT CONFLICT 90 90 S -> T (IN REF. 2). (POTENTIAL).
CC SQ SEQUENCE 345 AA; 37859 MW; 8C96PJDSCB0405 CRC64;
Query Match 33.2%; Score 613.5; DB 1; Length 345;
Best Local Similarity 38.2%; Pred. No. 3.9e-33;
Matches 138; Conservative 51; Mismatches 107; Indels 65; Gaps 10;
QY 15 ALLLGLVGLVSGL--SLEPVYVNSANKRFOAEGGYLYVPOIGDRDLICPRARPPPHS 72
DB 15 AMVVLTCRLATPLAKULBEPVSNLKPFLSGKGLVYIPKIGDKDITICPRAEGRP-- 72
QY 73 SPNTEFYKLYLVGAQRCEAPAPANNLLTCDRDLDFRTIKFOEYSPNLMGHEFRSH 132
DB 73 ---YEEYKLYLVPEQAAACSTVLDENVLTCKNPKHQEIRFTIKQEFSPNVMGIEFKKY 129
QY 133 HDVYITATSDGTREGESLGGVCLTRGMKVLIRVGQSRGCAVPRKPVSEMPERDRGA 192
DB 130 HDVYITSTNGSLGLEENREGVCRTRTKIWKVGQDP-NAVTPQLTTSRPSKESDNT 188
QY 193 AHSLEPGKENVLPDPTSNATSRGAGP-----LPPSPMAVAVGAAG-----LA 236
DB 189 VKT-----ATQAPRGSGQSDSGKHETVNOEKEKSGPAGAGGSGSDSDFNFK 236
QY 237 LLLLVGAAGGA-----MCMRRRAKPSSESHRPPGSGSGSLG-----GG 279
DB 237 VALFAAVGAGCVFLIIIFLTVLILKLRKRHKHTQ-----RAAALSTLSASP 287
QY 280 GCGMGREARPGELGALGGGAADPPCPHYEKVSDGHPVYTVODPSPSPNITYK 339
DB 288 KGGSGRAGTEPSDITILPLR--TTENNVCPEHYEKVSGDGHVYTVQEMPSPSPNITYK 344
QY 340 V 340
DB 345 V 345
RESULT 9
EFBI RAT
ID EFBI RAT STANDARD; PRT; 345 AA.
AC P52796;

DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ephrin-B1 precursor (EPH-related receptor tyrosine kinase ligand 2)
DE (LERK-2) (ELK ligand) (ELK-L).
GN EPHB1 OR EPHB2 OR LERK2.
OS Rattus norvegicus (Rat).
CC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC NCBI_TaxID=10116;
CC [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95022634; PubMed=7936648;
RA Fletcher P.A., Carpenter M., Shilling H., Baum P., Ziegler S.,
RA Gimpel S., Hollingworth T., Vandenbos T., Davison B.L.,
RA Lyman S.D., Beckmann M.P.,
RT "LERK-2, a binding protein for the receptor-tyrosine kinase Elk, is
RT evolutionarily conserved and expressed in a developmentally regulated
RT pattern";
RL Oncogene 9:3241-3248(1994).
CC -1- FUNCTION: BINDS TO, AND INDUCE THE COLLAPSE OF, COMMISSURAL
CC AXONS/GROWTH CONES IN VITRO. MAY PLAY A ROLE IN CONSTRAINING THE
CC ORIENTATION OF LONGITUDINALLY PROJECTING AXONS (BY SIMILARITY).
CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHB3 (PREFERRED),
CC EPHB1 AND EPHB2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- PTM: INDUCIBLE PHOSPHORYLATION OF TYROSINE RESIDUES IN THE
CC CYTOPLASMIC DOMAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U07560; AAA53092.1; -
CC InterPro: IPR001799; Ephrin.
CC Pfam: PF00812; Ephrin.1.
CC PRINTS: PRO1347; EPHRIN.
CC ProDom: PD002533; Ephrin.1.
CC PROSITE: PS01299; EPHRIN.1.
CC Developmental protein; Neurogenesis; Transmembrane; Glycoprotein;
CC Signal; Phosphorylation.
CC FT CHAIN 1 24 POTENTIAL.
CC FT DOMAIN 25 345 EPHRIN-B1.
CC FT TRANSMEM 237 257 EXTRACELLULAR (POTENTIAL).
CC FT DOMAIN 258 345 POTENTIAL.
CC FT DOMAIN 343 345 CYTOPLASMIC (POTENTIAL).
CC FT CARBOHYD 139 139 PDZ RECOGNITION MOTIF (POTENTIAL).
CC FT CONFLICT 90 90 N-LINKED (GLCNAC...) (POTENTIAL).
CC SQ SEQUENCE 345 AA; 37951 MW; 1B3045C5C73587E CRC64;
Query Match 32.9%; Score 608.5; DB 1; Length 345;
Best Local Similarity 38.0%; Pred. No. 8.2e-33;
Matches 137; Conservative 52; Mismatches 107; Indels 65; Gaps 10;
QY 15 ALLLGLVGLVSGL--SLEPVYVNSANKRFOAEGGYLYVPOIGDRDLICPRARPPPHS 72
DB 15 AMVVLTCRLATPLAKULBEPVSNLKPFLSGKGLVYIPKIGDKDITICPRAEGRP-- 72
QY 73 SPNTEFYKLYLVGAQRCEAPAPANNLLTCDRDLDFRTIKFOEYSPNLMGHEFRSH 132
DB 73 ---YEEYKLYLVPEQAAACSTVLDENVLTCKNPKHQEIRFTIKQEFSPNVMGIEFKKY 129
QY 133 HDVYITATSDGTREGESLGGVCLTRGMKVLIRVGQSRGCAVPRKPVSEMPERDRGA 192
DB 130 HDVYITSTNGSLGLEENREGVCRTRTKIWKVGQDP-NAVTPQLTTSRPSKESDNT 188
QY 193 AHSLEPGKENVLPDPTSNATSRGAGP-----LPPSPMAVAVGAAG-----LA 236

```

Db 189 YKT-----ATQAPGRSGQSDSDGKETVNOQEKSGPAGGSGSSDTSFNSK 236
QY 227 LLLLVAGAGGA-----MCMRRRAKPSERHRHGPSPFGSGSLG-----CG 279
Db 237 VALFAAVAGGCVIFLLIIIFLTVLLIKRKRRKHTQO-----DAALSLSTLASP 287
QY 280 GGGMGPREAPPEGLIALRGGAADPPFCPEHYEKVSGDGHPPYIVODGPSPENIYRK 339
Db 288 KQDSGFTAGTEPDIIIPLR---TENNVCPEHYEKVSGDGHPPYIVQEMPEQSPANIYRK 344
QY 340 V 340
Db 345 V 345

RESULT 10
EPB1_XENLA STANDARD; PRT; 327 AA.
ID EPB1_XENLA STANDARD; PRT; 327 AA.
AC 013097;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ephrin-B1 precursor (EPH-related receptor tyrosine kinase ligand 2) (LERK-2) (Elk ligand) (ELK-L) (XLERK).
GN EPB1 OR EPLG2 OR LERK2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_Taxid=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9716777; PubMed=9174051;
RA Jones T.L., Karavanova I., Chong L., Zhou R.P., Daar I.O.;
RT "Identification of Xlerk, an Eph family ligand regulated during mesoderm induction and neurogenesis in Xenopus laevis.";
RU Oncogene 14:2159-2166(1997).
CC -1- FUNCTION: MAY HAVE A ROLE IN THE DEVELOPING MESENCHYMAL AND NERVOUS TISSUE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: IN THE ADULT, EXPRESSED AT LOW LEVELS IN MOST ADULT TISSUES WITH INCREASED LEVELS OBSERVED IN THE KIDNEY, OOCYTES, OVARY AND TESTIS.
CC -1- PM: INDUCIBLE PHOSPHORYLATION OF TYROSINE RESIDUES IN THE CYTOPLASMIC DOMAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; U31427; AAC35995.1; -
CC InterPro: IPR001799; Ephrin.
CC Pfam: PF00812; Ephrin.1.
CC PRINTS; PR01347; EPHRIN.
CC ProDom; PD002533; Ephrin; 1.
CC PROSITE; PS01299; EPHRIN; 1.
CC Developmental protein; Neurogenesis; Transmembrane; Glycoprotein;
CC Signal; Phosphorylation.
CC CHAIN 1 20 POTENTIAL.
CC FT CHAIN 21 225 EPHRIN-B1.
CC FT DOMAIN 21 225 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 226 246 POTENTIAL.
CC FT DOMAIN 247 327 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 325 327 PDZ RECOGNITION MOTIF (POTENTIAL).
CC CAROHD 131 131 N-LINKED (GLCNAc. . .) (POTENTIAL).
CC FT CAROHD 202 202 N-LINKED (GLCNAc. . .) (POTENTIAL).
CC SQ SEQUENCE 327 AA; 36621 MW; 71230CE7F6B5974 CRC64;

```

```

Query Match 32.4%; Score 600; DB 1; Length 327;
Best Local Similarity 39.8%; Pred. No. 2,8e-32;
Matches 146; Conservative 43; Mismatches 100; Indels 78; Gaps 12;

QY 10 GVR--VQALLIGVGLVSGLSLEPYWNSANKRQAEGLYVYIQIGRLDLCRA-- 65
Db 3 GRRRLGLLVYIRLCSAAGKNLEPYWNSQNPRTISGVLVYELIGRDLIIICPKGLF 62
QY 66 RPPGPHSPSEYEFYKLYLVGAQGRCEAPNLLLTCDRPPDLIRFTIKFOEYSPNL 125
Db 63 QP-----YETKLYMRDRQLEACSTVDPNVLYVTONQGEKRFITIKQESPNM 114
QY 126 GHEFRSHDYIYIATSDGTREGLESLOGVCVLTGKVKLLRVGSGPRGAVRKPYSKP 185
Db 115 GLEFRNDQYIYIATSDGTREGLESLOGVCVLTGKVKLLRVGSGPRGAVRKPYSKP 169
QY 186 MERDRGAHSLERGENLGDPTSNATSRGA-EGPLPFP-----SMPAVAGAAGLA 236
Db 170 -----TTPRSK-----ADNTGKIATGPNNGVQNGKSDTMSDKPTGRWGVDFE 217
QY 237 LLLLVAGAGGAMC-----MRRRAKPS-----SRHPGSGFGRG 273
Db 218 NSKIATFAAIGAGCVIFLLIIIFLVVLLIKRKRRKHTQQAALSLSTLSPKSGNAG 277
QY 274 SLGSGGGGNGPREAPPEGLIALRGGAADPPFCPEHYEKVSGDGHPPYIVODGPSP 333
Db 278 S-----EPDIIIPLR---TENNVCPEHYEKVSGDGHPPYIVQEMPEQSP 320
QY 334 PNYYKY 340
Db 321 ANIYKY 327

RESULT 11
EPB2_BRARE STANDARD; PRT; 195 AA.
ID EPB2_BRARE STANDARD; PRT; 195 AA.
AC P79727;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ephrin-A2 precursor (EPH-related receptor tyrosine kinase ligand 6) (LERK-6) (ELF-1) (ZEPHL3).
GN EPB2 OR EPLG6 OR LERK6.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_Taxid=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97195707; PubMed=9043080;
RA Brennan C., Monschau B., Lindberg R., Guthrie B., Dreacher U.,
RA Bonhoeffer F., Holder N.;
RT "Two Eph receptor tyrosine kinase ligands control axon growth and may be involved in the creation of the retinotectal map in the zebrafish.";
RU Development 124:655-664(1997).
CC -1- FUNCTION: CONTROL AXON GROWTH AND MAY BE INVOLVED IN THE CREATION OF THE RETINO-TECTAL MAP.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (Potential).
CC -1- TISSUE SPECIFICITY: WIDESPREAD EXPRESSION IN THE EMBRYO.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE PRESUMPTIVE MIDBRAIN OF DEVELOPING EMBRYOS FROM THE SIX-SOMITE STAGE. BY 24 HOURS IT IS EXPRESSED THROUGHOUT THE MIDBRAIN INCLUDING THE REGION OF THE PRESUMPTIVE TECTUM. AT LATER STAGES IT IS EXPRESSED IN A GRADED FASHION THROUGHOUT THE TECTUM.
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y09668; CAA07863.1; -
 DR ZFIN; ZDB-GENE-990415-66; etna2.
 DR InterPro; IPR001799; Ephrin.
 DR Pfam; PF00812; Ephrin; 1.
 DR PRINTS; PR01347; EPHRIN.
 DR ProDom; PD002533; Ephrin; 1.
 DR PROSITE; PS01299; EPHRIN; 1.
 DR Developmental protein; Neurogenesis; Glycoprotein; GPI-anchor; Signal.
 KW SIGNAL
 FT SIGNAL 1 16 POTENTIAL.
 FT CHAIN 17 195 EPHRIN-A2.
 FT CARBOHYD 32 32 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 195 AA; 22688 MW; 9EE284FEB61D0C42 CRC64;
 Query Match 10.0%; Score 185; DB 1; Length 195;
 Best Local Similarity 29.9%; Pred. No. 1.2e-05;
 Matches 63; Conservative 21; Mismatches 71; Indels 56; Gaps 7;
 QY 33 VYVNSANKFQAEAGYVLPOLGRLDLCPPRPPGPHSPNYEFYKLYVGAGQGRRC 92
 DB 29 VYVNSNSNRW-QGEYTVASINDYDVCPYTESPQHS--RMERYILFMVNHGYLETC 85
 QY 93 EAPPAENLLTCR---PDILRFITKFOEYSPNLMGHEFRSHDYIATSDGTREGLE 149
 DB 86 EHRMGEPKMECKRPPSPGPIRFSKFLFTEPFSIGFEFRDHEYYIYSSHPHNAKXP 145
 QY 150 SLGGVCLTRGMVTLRVQSPRGAVPKPYSEMERDRGAHSLBEGKXENLPGDPTS 209
 DB 146 CLK-----LKVYV-----KPYSS-----GYSEPPPLTD 169
 QY 210 NATSRGAEGPLPPSPAVAGAGLALLL 240
 DB 170 QSGRCAGDGPC-----LAVML 186
 RESULT 12
 EFA3_HUMAN STANDARD; PRT; 238 AA.
 ID EFA3_HUMAN
 AC P52797;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-UTL-1999 (Rel. 38, Last annotation update)
 DE Ephrin-A3 precursor (EPH-related receptor tyrosine kinase ligand 3)
 DE (IERK-3) (EHK1 ligand) (EHK1-L).
 GN EFN3 OR EPLG3 OR IERK3 OR EFL-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCB1_Taxid=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95140419; PubMed=7838529;
 RA Kozlosky C.J., Maraskovsky E., McGraw J.T., Vandenbos T.,
 RA Teepe M., Lyman S.D., Srinivasan S., Fletcher F.A., Gayle R.B. III,
 RA Cerretti D.P., Beckmann M.P.,
 RT "Ligands for the receptor tyrosine kinases hek and elk: isolation of
 RT cDNA encoding a family of proteins."
 RL Oncogene 10:299-306(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95063919; PubMed=7973638;
 RA Davis S., Gale N.W., Aldrich T.H., Maisonnier P.C., Uhotak V.,
 RA Pawson T., Goldfarb W., Yancopoulos G.D.;
 RT "Ligands for EPH-related receptor tyrosine kinases that require
 RT membrane attachment or clustering for activity."
 RL Science 266:816-819(1994).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, SKELETAL MUSCLE, SPLEEN,

CC THYMUS, PROSTATE, TESTIS, OVARY, SMALL INTESTINE, AND PERIPHERAL
 CC BLOOD LEUKOCYTES.
 CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U14187; AAC50078.1; -
 DR EMBL; L37360; AAA52368.1; -
 DR Genew; HGNC:3223; EFN3.
 DR MIM; 601381; -
 DR InterPro; IPR001799; Ephrin.
 DR Pfam; PF00812; Ephrin; 1.
 DR PRINTS; PR01347; EPHRIN.
 DR ProDom; PD002533; Ephrin; 1.
 DR PROSITE; PS01299; EPHRIN; 1.
 DR Glycoprotein; GPI-anchor; Signal.
 KW SIGNAL
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 238 EPHRIN-A3.
 FT CARBOHYD 38 38 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 67 67 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 100 100 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 71 74 MISSING (IN REF. 2).
 SQ SEQUENCE 238 AA; 26350 MW; 8EFD6A8FE33FDDA CRC64;
 Query Match 9.7%; Score 179; DB 1; Length 238;
 Best Local Similarity 28.4%; Pred. No. 3.6e-05;
 Matches 65; Conservative 24; Mismatches 80; Indels 60; Gaps 12;
 QY 7 GPGGVVAGALLLVGLVGLSLPYVNSANKRQAEAGYVLPOLGRLDLCR-R 64
 DB 24 GPG-----GALG-----NRHAVVNSNQHLRE-GYTVQVAVNDYDIYCPHYN 67
 QY 65 ARPPGPHSP---NVEFYKLYVGAGQGRCEAPRPNLLTCRPRDL---DLRFITK 117
 DB 68 SSGVGPAGPDPGGGAGQYVLYVNSNGYRTCNASQGR-RMECNRHAPHSPKSEK 126
 QY 118 QEYSPNLMGHEFRSHDYIATSDGTREGLESLOGVCLTRGMVTLRVQSPRGA 177
 DB 127 QRYGAFSLGYEFHAGHYIYISTPTNHLH-----WKCLR--MKVVCASHSHG--- 174
 QY 178 RKPVSERP-----MERDRGAHSLF-----PKENLP 204
 DB 175 EKPVTLPOFTMGPNVXINVLDEFGENPQVPLKESISGTSKREHL 223
 RESULT 13
 EFA2_MOUSE STANDARD; PRT; 209 AA.
 ID EFA2_MOUSE
 AC P52801;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Ephrin-A2 precursor (EPH-related receptor tyrosine kinase ligand 6)
 DE (IERK-6) (ELF-1) (CEK7-ligand) (CEK7-L).
 GN EFN2 OR EPLG6 OR IERK6 OR ELF1 OR EPL6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCB1_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=Swiss Webster; TISSUE=Brain;
 RA MEDLINE=95007776; PubMed=7522971;
 RA Cheng H.J., Flanagan J.G.;
 RT "Identification and cloning of ELF-1, a developmentally expressed
 RT ligand for the Mekk and sek receptor tyrosine kinases."
 RL Cell 79:157-168(1994).

```

[2]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RX MEDLINE=95181289; PubMed=7876076;
RA Shao H., Lou L., Pandey A., Verdier M.F., Siever D.A., Dixit V.K.;
RT "cDNA cloning and characterization of a Cx47 receptor
RT protein-tyrosine kinase ligand that is identical to the ligand
RT (SLF-1) for the Mek-4 and Ser receptor protein-tyrosine kinases";
RL J. Biol. Chem. 270:3467-3470(1995).
CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA3, EPHA4 AND
CC EPHA5.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U14941; AAA53636.1; -
DR EMBL: U14752; AAA65520.1; -
DR MGD; MG1:102707; Etna2.
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin; 1.
DR PRINTS; PR01347; EPHRIN.
DR ProDom; PD002533; Ephrin; 1.
DR PROSITE; PS01299; EPHRIN; 1.
KM Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 20
FT CHAIN 21 209
FT EPHRIN-A2.
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 170 170 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 209 AA; 23586 MW; P197545F25B9ABC CRC64;

Query Match 9.5%; Score 176; DB 1; Length 209;
Best Local Similarity 29.3%; Pred. No. 4.9e-05;
Matches 58; Conservative 19; Mismatches 69; Indels 52; Gaps 7;

OY 33 VYMNANKRFOA-----GGVLYPQIGDRDLCPARPFGPHSSNVEFYKLYVGA 87
DB 35 VYMNANKRFOVASVAGDGGGTVVEVSINDYLDYCPHYGALP-PARMEHYVLYMVNGE 93
OY 88 QGRCEAPAPNLLITCDRPL--D-RFTIKFOEYSPNLMGHEFRSHDYIATSDGT 144
DB 94 GHASCDHRGQFKWECNRPAPGAPLKFSEKFLTFPFSLGFEFRGHEYYISATP-- 151
OY 145 REGESLOGGVCITRGMKYLIRVGSPRGAVPRKPSSEMERDRGAASLEPKENLP 204
DB 152 ---PNLVDRPCLR-LKYYVR-----PTNETLY 174
OY 205 GDP-----TSNATSRGAGS 218
DB 175 EAPERTFTSNSSCSGLGS 192

RESULT 14
EFA2 HUMAN STANDARD; PRT; 213 AA.
AC 043921; 076020;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ephrin-A2 precursor (EPH-related receptor tyrosine kinase ligand 6)
DE (LRRK-6) (HEK7-ligand) (HEK7-L).
GN EFA2 OR EPLG6 OR LRRK6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumetazoa;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98126446; PubMed=9465306;
RA Cerretti D.P., Nelson N.;
RT "Characterization of the genes for mouse LRRK-3/Ephrin-A3 (Epl3)
RT mouse LRRK-4/Ephrin-A4 (Epl4), and human LRRK-6/Ephrin-A2 (EPLG6);
RT conservation of intron/exon structure.";
RL Genomics 47:131-135(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,
RA Burkhart-Schultz K., Gordon L., Kyle A., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Garner J., Danganan L., Poundstone P.,
RA Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andreise T.,
RA Traubheim M., Amico-Keller G., Coefield J., Dharie S., Lucas S.,
RA Bruce R., Thomas P., Quan G., Kronmiller B., Arellano A.,
RA Montgomey M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.O.,
RA Carraro A.V.;
RA Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=99045414; PubMed=9826538;
RA Aashelm H.C., Pedetour F., Grosgeorge J., Logtenberg T.;
RT "Cloning, chromosomal mapping, and tissue expression of the gene
RT encoding the human Eph-family kinase ligand ephrin-A2.";
RL Biochem. Biophys. Res. Commun. 252:378-382(1998).
CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA3, EPHA4 AND
CC EPHA5.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U92896; AAC39577.1; -
DR EMBL: U92893; AAC39577.1; JOINED.
DR EMBL: U92894; AAC39577.1; JOINED.
DR EMBL: AC004286; AAC04836.1; -
DR EMBL: AC004292; CA007435.1; -
DR GeneW; HGNC:3222; EPHA2.
DR MIM; 602756; -
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin; 1.
DR PRINTS; PR01347; EPHRIN.
DR ProDom; PD002533; Ephrin; 1.
DR PROSITE; PS01299; EPHRIN; 1.
KM Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 24
FT CHAIN 25 213
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 6 6 R->A (IN REF. 3).
FT CONFLICT 25 26 R->P (IN REF. 3).
FT CONFLICT 29 30 AA->RR (IN REF. 3).
SQ SEQUENCE 213 AA; 23678 MW; 33C9FB1A8168B2D0 CRC64;

Query Match 9.5%; Score 175.5; DB 1; Length 213;
Best Local Similarity 36.8%; Pred. No. 5.4e-05;
Matches 43; Conservative 14; Mismatches 51; Indels 9; Gaps 3;

OY 33 VYMNANKRFOA-----GGVLYPQIGDRDLCPARPFGPHSSNVEFYKLYVGA 87
DB 39 VYMNANKRFOAGDADGGGTVVEVSINDYLDYCPHYGALP-PARMEHYVLYMVNGE 97

```

QY 88 QGRRCCEAPPANLTLTCDRDL---DIRFTIKQEYSPNMGHEPSSHDDYTIATS 141
DB 98 GHASCDHRORGFKEWECNRPAAPGPKFSEBKQFTPLFELGFEFRPGHEYYIATS 154

RESULT 15

EPF2_CHICK STANDARD; PRT; 200 AA.

ID EPF2_CHICK
AC P52802;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ephrin-A2 precursor (EPH-related receptor tyrosine kinase ligand 6)
DE (ERRK-6) (ELF-1).
GN EFNA2 OR EPLG6 OR LERK6 OR ELF1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95360981; PubMed=7634327;
RA Cheng H.J., Nakamoto M., Bergmann A.D., Flanagan J.G.;
RT "Complementary gradients in expression and binding of ELF-1 and Mek4
in development of the topographic retinorectal projection map";
RL Cell 82:371-381(1995).
CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA3, EPHA4 AND
EPAH5 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
(potential).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN A GRADIENT ACROSS THE TECTUM
BEING MORE STRONGLY EXPRESSED AT THE POSTERIOR POLE.
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; LA0933; AAC42229.1; -
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin.1.
DR PRINTS; PR01347; EPHRIN.
DR ProDom; PD002533; Ephrin.1.
DR PROSITE; PS01299; EPHRIN.1.
KW Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 22
FT CHAIN 1 22
FT CARBOHYD 23 36
FT CARBOHYD 161 161
FT CARBOHYD 175 175
SQ SEQUENCE 200 AA; 23049 MW; 8FAD1AE5E45ED96 CRC64;

Query Match 9.3%; Score 172; DB 1; Length 200;

Best Local Similarity 35.3%; Pred. No. 8.5e-05; Indels 16; Gaps 5;
Matches 49; Conservative 16; Mismatches 58;

QY 15 ALLILGLVGLVSGISLEP-----VYNSANKRFOAGGVLYPQIGDRDLCPRA 65
DB 7 AALLAIVG-VCVMSDDPGKVISDRYAVVYNNRGNPRFH-RGDTYVEVSINDYLDICPHY 64
QY 66 RPPGPHSPHYEYKYLYVGAQGRCEAPPANLTLTCDRDL---DIRFTIKQEYSP 122
DB 65 EEPPLP-AEMEMEVLYLWYVYEGHASCDRKQKFKWECNRPDSPGPKFSEBKQFTPL 122
QY 123 NWGHEFRSHDDYTIATS 141
DB 123 FSLGFEFRPGHEYYIATS 141

RESULT 16

EPF5_CHICK STANDARD; PRT; 228 AA.

ID EPF5_CHICK
AC P52804;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ephrin-A5 precursor (EPH-related receptor tyrosine kinase ligand 7)
DE (ERRK-7) (RAGS protein).
GN EFNA5 OR RAGS (Chicken).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Posterior tectum;
RX MEDLINE=95360980; PubMed=7634326;
RA Drescher U., Kremoser C., Handwerker C., Loschinger J., Noda M.,
RA Bonhoeffer F.;
RT "In vitro guidance of retinal ganglion cell axons by RAGS, a 25 kDa
tectal protein related to ligands for Eph receptor tyrosine
kinases";
RL Cell 82:359-370(1995).
CC -1- FUNCTION: INDUCES GROWTH CONE COLLAPSE AND REPUSION OF RETINAL
GANGLION CELL AXONS.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
(potential).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN A GRADED FASHION ACROSS THE
TECTUM BEING MORE STRONGLY EXPRESSED TOWARDS THE POSTERIOR POLE.
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X90377; CA62027.1; -
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin.1.
DR PRINTS; PR01347; EPHRIN.
DR ProDom; PD002533; Ephrin.1.
DR PROSITE; PS01299; EPHRIN.1.
KW Developmental protein; Neurogenesis; Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 20
FT CHAIN 1 20
FT CARBOHYD 21 228
FT CARBOHYD 37 37
SQ SEQUENCE 228 AA; 26206 MW; 56DB84FBDECF18AD CRC64;

Query Match 9.2%; Score 170.5; DB 1; Length 228;

Best Local Similarity 28.9%; Pred. No. 0.00012; Indels 59; Gaps 13;
Matches 73; Conservative 30; Mismatches 91;

QY 16 LLLILGLVGLVSGISLEP-----VYNSANKRFOAGGVLYPQIGDRDLCPRA 64
DB 6 MLLIAVALWVCGVGRGPRKAVADRYAVYNNSTNPRFQ-QGDHIVDCINDYLDVCPH 64
QY 65 ARPGRPHSPHYEYKYLYVGAQGRCEAPPANLTLTCDRDLCPRA 112
DB 65 YEDVEPDKT--ERYVLYWVNFQYSSCHISNGFKWECNRPHSPV-----GPKL 113
QY 113 FTTFQEYSPNMGHEFRSHDDYTIATS---DGTREGLESLOGGVLTGMKYLAVGQ 169
DB 114 FSEKFOFTPLPSLGFEPFRGEXEYFYSALPDNGRRS-----CLK-LKVFV--- 159
QY 170 SPRGAVPRKPVSEMPERDCAHSLPEKENIPGPTSNATSRGAGCLPSPSP--A 227
DB 160 -PANSCKMTIGVHDVRVDVADKVENSLPADRV---RESAEPGRG-ENNAQTPRIPIRL 214

QY 228 VAGAAGLALLL 240
 ID 215 LATTLLPLAMLLI 227

RESULT 17
 EFAS_HUMAN STANDARD; PRT; 201 AA.

AC P52798; O95457; Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ephrin-A4 precursor (EPH-related receptor tyrosine kinase ligand 4)
 DE (LEK-4)
 GN EPHRA4 OR EPLG4 OR LEK4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 ON NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=95140419; PubMed=783529;
 RA Kozlovsky C.J., Maraskovsky E., McGrew J.T., Vandenbos T.,
 RA Teepe M., Lyman S.D., Srinivasan S., Fletcher F.A., Gayle R.B. III,
 RA Cerretti D.P., Beckmann M.P.;
 RT "ligands for the receptor tyrosine kinases hek and elk: isolation of
 RT cDNAs encoding a family of proteins";
 RL Oncogene 10:1239-306(1995).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX TISSUE=B-cell;
 RX MEDLINE=20076261; PubMed=10607706;
 RA Aasheim H.C., Munthe E., Funderud S., Smeland E.B., Betske K.,
 RA Logtenberg T.;
 RT "A splice variant of human ephrin-A4 encodes a soluble molecule that
 RT is secreted by activated human B lymphocytes.";
 RL Blood 95:1221-230(2000).
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE INTERACTION BETWEEN ACTIVATED B
 CC LYMPHOCYTES AND DENDRITIC CELLS IN TONSILS.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
 CC OR SECRETED (DEPENDING ON THE ISOFORM).
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1/GPI-anchored (shown here) and
 CC 2/secreted; are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE ADULT SPLEEN, LYMPH NODE,
 CC PROSTATE, OVARY, SMALL INTESTINE, AND COLON, AND IN PERIL HEART.
 CC LUNG, LIVER AND KIDNEY. ALSO DETECTED IN HEPA200 EPITHELIAL CELL LINES.
 CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U14188; AAC50079.1; -
 DR EMBL; AJ006352; CAA06992.1; -
 DR EMBL; AJ006353; CAA06993.1; -
 DR Genew; HGNC:3224; EPHNA4.
 DR MTW; 601380; -
 DR InterPro; IPR001799; Ephrin.
 DR Pfam; PF00812; Ephrin.1.
 DR PRINTS; PRO1347; EPHRIN.
 DR ProDom; PD002533; Ephrin.1.
 DR PROSITE; PS01299; EPHRIN.1.
 KM Glycoprotein; GPI-anchor; Signal; Alternative splicing.
 FT SIGNAL 1 22
 FT CHAIN 1 170
 FT PROPEP 171 201
 FT CARBOHYD 33 33
 FT LIPID 170 170
 FT REMOVED IN MATURE FORM (POTENTIAL).
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 FT GPI-ANCHOR (POTENTIAL).

FT VARSPLIC 157 201 KSSAHFVGSFGESGTCGRCGDTPPSLCLLLLLILRL
 FT LRL -> NLPFSHKEPSSSQDPLEEGSLPLALGVPIQTD
 FT KME (IN ISOFORM 2).
 SQ SEQUENCE 201 AA; 22386 MM; ABE8D5443A9AF28D CRC64;

Query Match 9.2%; Score 169.5; DB 1; Length 201;
 Best Local Similarity 29.9%; Pred. No. 0.00012;
 Matches 66; Conservative 18; Mismatches 82; Indels 55; Gaps 10;

QY 24 LVSGLSL-EPYVNSANKRFQAEGLVLPQIGRLDLCRPAPPGHSPNTEFYLY 82
 DB 20 LRGSLSLHHVYVYVNSNRL-LRGDAVELGLINDVIDICPHYEGPPGEP--ETFAly 76

QY 83 LVGAQGRCEAP-PAPVLLTCDRPDLARPTIKFQVSYNLMGHERSHHDYLLITS 141
 DB 77 VDWDPGYSQAGPRAYKRVCSLPFGHVPSKIQRTFPLSLGFELPGERTYIYISVP 136

QY 142 DGTREGLSLQGGVCLTRGMKVLRVGSPFGAVPRKPVSEMPERDRGAHSLPGEKE 201
 DB 137 --TPE-----SSGQCL-----RLQGVCCCKERSGSAHPV----- 164

QY 202 NLPDPTSNATS--RGAGPLPPEPMFVAGAAGLALLL 240
 DB 165 --GSPGESGTSQWEGADPTPSP-----LCULLL 189

RESULT 18
 EFAS_BRARE STANDARD; PRT; 228 AA.

ID EFAS_BRARE
 AC P79728;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ephrin-A5 precursor (EPH-related receptor tyrosine kinase ligand 7)
 DE (LEK-7) (AL-1) (ZEPHRA4)
 GN EFNA5 OR EFNA5B OR AL1 OR EPLG7 OR LEK7.
 OS Brachyotia retilio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 ON NCBI_TaxID=7955;
 [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=embryo;
 RX MEDLINE=97195707; PubMed=9043080;
 RA Brennan C., Monseu B., Lindberg R., Guthrie B., Drescher U.,
 RA Bonhoeffer F., Holder N.;
 RT "Two Ebn receptor tyrosine kinase ligands control axon growth and may
 RT be involved in the creation of the retinotectal map in the
 RT zebrafish";
 RL Development 124:655-664(1997).
 CC -1- FUNCTION: CONTROL AXON GROWTH AND MAY BE INVOLVED IN THE CREATION
 CC OF THE RETINO-TECTAL MAP.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (potential).
 CC -1- TISSUE SPECIFICITY: WIDESPREAD EXPRESSION IN THE EMBRYO.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE PRESUMPTIVE MIDBRAIN OF
 CC DEVELOPING EMBRYOS FROM THE SIX-SOMITE STAGE. BY 24 HOURS IT IS
 CC STRONGLY EXPRESSED IN THE MIDBRAIN CAUDAL TO THE PRESUMPTIVE
 CC TECTUM. AT LATER STAGES IT IS MAINTAINED AT THE POSTERIOR MARGIN
 CC OF THE TECTUM.
 CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y09669; CAA70864.1; -
 DR ZFIN; ZDB-GENE-980526-186; efna5b.

DR InterPro: IPR001799; Ephrin.
 DR Pfam: PF00812; Ephrin.1.
 DR PRINTS: PR01347; EPHRIN.1.
 DR PRODOM: PD002533; Ephrin.1.
 DR PROSITE: PS01299; EPHRIN.1.
 DR Developmental protein; Neurogenesis; Glycoprotein; GPI-anchor; Signal.
 FT SIGNAL 1 20
 FT CHAIN 1 228
 FT CARBOHYD 37 37
 FT CARBOHYD 37 37
 SQ SEQUENCE 228 AA; 26595 MW; 74B3406C05418E6E CRC64; (POTENTIAL).
 Query Match
 Best Local Similarity 27.6%; Pred. No. 0.00014;
 Matches 72; Conservative 28; Mismatches 84; Indels 77; Gaps 13;
 9.2%; Score 169.5; DB 1; Length 228;
 16 LILLGLGVLSGLSP-----VYNSANKRFOAGGVLYPQIGRLDLLCPRA 65
 7 IYFVGVLWVCVPSQSPSSKVNADRYAVFWMRTNPRFQ-RGDHIDVCINDYIDVCPHY 65
 66 RPPGPHSSPNVEFYKLYVG-----GAQGRCEAPAPNLLTCDPDLRF 113
 66 EDVPEE-RTERVYLYMNYDGYSTCDHTAGFKMECNRPSPV-----GPIKF 114
 114 TIFQYSPMLNCHERSHDYITIA---TSQTRGSLQGVCLTRGMKYLIRVQS 170
 115 SEKFQUTFPFSLGFERRPQREYYIISMTETGRS-----CLK-LKVFYR---P 160
 171 PRG-----GAVPRKPYSEM---PMERDRGAASLBPCKENLPQDPSTNATSGAEP 222
 161 PNCCEKTIQVHDRVFDKVDNALBPRDTSHEAPRSQVSTSGLRKQTSR---PL-- 214
 223 PSMFAVGAAGLALLLGA 243
 215 -----LALLLCIS 223
 RESULT 19
 EFAL_XENIA
 ID EFAL_XENIA STANDARD; PRT; 216 AA.
 AC P52794;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ephrin-A1 precursor (EPH-related receptor tyrosine kinase ligand 1)
 GN EFN1 OR EPG1 OR ELF.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodidae; Xenopus.
 NC NCB1_TaxID=8355;
 RN NCB1_TaxID=8355;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97000306; PubMed=8843391;
 RA Weinstein D.C., Rahman S.M., Ruiz J.C., Hemmati-Brivanlou A.;
 RT "Embryonic expression of eph signaling factors in Xenopus.";
 RL Mech. Dev. 57:113-144 (1996).
 CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA2, EPHA4,
 CC EPHA5, EPHA6 AND EPHA7. ALSO BINDS WITH LOW AFFINITY TO EPHA1 (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (BY
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; U31204; AAA74485.1; -

DR EMBL: U31205; AAA74486.1; -
 DR InterPro: IPR001799; Ephrin.
 DR Pfam: PF00812; Ephrin.1.
 DR PRINTS: PR01347; EPHRIN.1.
 DR PRODOM: PD002533; Ephrin.1.
 DR PROSITE: PS01299; EPHRIN.1.
 DR Glycoprotein; GPI-anchor; Alternative splicing; Signal.
 FT SIGNAL 1 28
 FT CHAIN 1 228
 FT CARBOHYD 36 36
 FT CARBOHYD 36 36
 FT VARSPLIC 162 216
 SQ SEQUENCE 216 AA; 24755 MW; 1B3A508E0A7B872E CRC64;
 Query Match
 Best Local Similarity 29.2%; Pred. No. 0.00018;
 Matches 71; Conservative 26; Mismatches 93; Indels 53; Gaps 12;
 9.1%; Score 167.5; DB 1; Length 216;
 12 RVGALLLVGLVSGLSLEP-----VYNSANKRFOAGGVLYPQIGRLDLLCPRA 64
 6 RAAVQLIVGV-GLGVGLMLRBAQGERHIVFMSNTRFMQE-DYTVQVQLNDYIDVCPY 63
 65 ARPPGPHSSPNVEFYKLYVGAGQGRCEAPAPNLLTCDPDL--DLPFTIKQYIS 121
 64 YE-BGSVAQHTVERKYLFLVDEYETCKRSDQVMECNKFPAPGPKFCKQKFT 122
 122 PNLGHEFRSHDYYITATSDGTREGLESLOGGVCLTRGMKYLIRVQSPRGAVPRKPY 181
 123 PFTGTGTFEPRERYIYSK-----IHYHGTCNR-----LRHVS---GRTPPEV 166
 182 SEMPERDRGAASLBPCKENLPQDPSTNATSGAEPPLPSPMPAVAG--AAGALLL 229
 167 N-----VHT-----PSHSIQSDPEVPL-PGVKSVAGNSAAPGIPCTL 204
 240 LGV 242
 205 YGL 207
 RESULT 20
 EFAS_HUMAN
 ID EFAS_HUMAN STANDARD; PRT; 228 AA.
 AC P52803;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ephrin-A5 precursor (EPH-related receptor tyrosine kinase ligand 7)
 GN EFN5 OR EPIG7 OR LERK7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC NCB1_TaxID=9606;
 RN NCB1_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95267434; PubMed=7748564;
 RA Winslow J.W., Moran P., Valverde J., Shih A., Yuan J.Q., Wong S.C.,
 RT "ERK-7: a ligand of the Eph-related kinases is developmentally
 RT regulated in the brain.";
 RL Cytokine 9:540-549 (1997).
 CC [3]
 CC FUNCTION.
 CC MEDLINE=20069483; PubMed=10601038;
 RA Davy A., Gale N.W., Murray E.W., Klinghoffer R.A., Soriano P.,

RA Feuerstein C., Robbins S.M.;
 RT "Compartmentalized signaling by GPI-anchored ephrin-A5 requires the
 RT Fyn tyrosine kinase to regulate cellular adhesion."
 RL Genes Dev. 13:3125-3135 (1999).
 CC -1- FUNCTION: MAY FUNCTION ACTIVELY TO STIMULATE AXON FASCICULATION.
 CC INDUCES COMPARTMENTALIZED SIGNALING WITHIN A CAVEOLAE-LIKE
 CC MEMBRANE MICRODOMAIN WHEN BOUND TO THE EXTRACELLULAR DOMAIN OF ITS
 CC COGNATE RECEPTOR. THIS SIGNALING EVENT REQUIRES THE ACTIVITY OF
 CC THE FYN TYROSINE KINASE.
 CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA2, EPHA3 AND
 CC EPHB1.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR IT
 CC IS COMPARTMENTALIZED IN DISCRETE CAVEOLAE-LIKE MEMBRANE
 CC MICRODOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U26403; AAB60377.1; -.
 CC Genew; HGNC:3225; EFNA5.
 CC MIM; 601535; -.
 CC InterPro: IPR001799; Ephrin.
 CC Pfam; PF00812; Ephrin.1.
 CC PRINTS; PR01347; EPHRIN.
 CC ProDom; PD002533; Ephrin.1.
 CC PROSITE; PS01299; EPHRIN.1.
 CC Developmental protein; Neurogenesis; Glycoprotein; GPI-anchor; signal;
 CC Polymorphism.
 CC FT SIGNAL 1 20 POTENTIAL.
 CC CHAIN 21 228 EPHRIN-A5.
 CC CARBOHYD 37 37 N-LINKED (GLCNAC...) (POTENTIAL).
 CC VARIANT 55 55 N->K (IN DBSNP:469062).
 CC FT VARSLIC /FTID=VAR_012035.
 CC SQ SEQUENCE 228 AA; 26297 MW; 6893B1CCACFF3F57 CRC64;
 CC -----
 CC Query Match 9.1%; Score 167.5; DB 1; Length 228;
 CC Best Local Similarity 28.8%; Pred. No. 0.00019;
 CC Matches 65; Conservative 29; Mismatches 81; Indels 51; Gaps 11;
 CC -----
 CC QY 33 VYVNSANKRFOAEAGYVLPQIGRLDILCPAPRPPGSHSSPNVEFYKLYLVG----- 85
 CC DB 34 VYVNSNSNPRFO-RGDYHIDVLCINDYLVCPHYEDSVPEDKT--ERYVLVYVNFQGSAC 90
 CC QY 86 -----GAQGRRCAPAPNLLTCORPDLILFTTKFOYSFNLWGHFRSHDYIYIAT 140
 CC DB 91 DHTSKGFKEWCKNRHPSN-----GPIKFSEKFOLETFPSLGFRRGRREYFYIS 141
 CC QY 141 S---DGTREGLESLQGVCLTRGMKYLRLVQGSPRGAVPRKPVSEMERDRGAASLE 197
 CC DB 142 AIPDNGRRS-----CLK--LKVFVR-----PTNSCKMTIGVDRVVDVNDKVENSL 186
 CC QY 198 PKENLPDPTSNATSRGAEGPLPPSPMPAVAGAAGLALLILGVA 243
 CC DB 187 PADDTV---HESAPSRG-ENNAQTPTIPSRL-----LAILFLLA 223
 CC -----
 CC RESULT 21
 CC EFNA5_MOUSE STANDARD; PRT; 228 AA.
 CC AC 008543; 008544;
 CC DT 01-NOV-1997 (Rel. 35; Created)
 CC DT 01-NOV-1997 (Rel. 35; Last sequence update)
 CC DT 30-MAY-2000 (Rel. 39; Last annotation update)
 CC DE Ephrin-A5 precursor (Eph-related receptor tyrosine kinase ligand 7)
 CC (LEK-7) (AL-1).
 CC GN EFNA5 OR EPLG7 OR LEK7 OR EPL7.
 CC OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muride; Murinae; Mus.
 CC NCBI_taxid=10090;
 CC RN (1)
 CC SEQUENCE FROM N.A.
 CC MEDLINE=97060319; PubMed=8903354;
 CC RA Flenikien A.M., Gale N.W., Yancopoulos G.D., Wilkinson D.G.;
 CC "Distinct and overlapping expression patterns of ligands for
 CC Eph-related receptor tyrosine kinases during mouse embryogenesis";
 CC Dev. Biol. 179:382-401 (1996).
 CC -1- FUNCTION: MAY FUNCTION ACTIVELY TO STIMULATE AXON FASCICULATION.
 CC INDUCES COMPARTMENTALIZED SIGNALING WITHIN A CAVEOLAE-LIKE
 CC MEMBRANE MICRODOMAIN WHEN BOUND TO THE EXTRACELLULAR DOMAIN OF ITS
 CC COGNATE RECEPTOR. THIS SIGNALING EVENT REQUIRES THE ACTIVITY OF
 CC THE FYN TYROSINE KINASE (BY SIMILARITY).
 CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA2, EPHA3 AND
 CC EPHB1 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR IT
 CC IS COMPARTMENTALIZED IN DISCRETE CAVEOLAE-LIKE MEMBRANE
 CC MICRODOMAINS (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U90664; AAB50239.1; -.
 CC EMBL; U90665; AAB50240.1; -.
 CC MGD; MG1:107444; Efna5.
 CC InterPro: IPR001799; Ephrin.
 CC Pfam; PF00812; Ephrin.1.
 CC PRINTS; PR01347; EPHRIN.
 CC ProDom; PD002533; Ephrin.1.
 CC PROSITE; PS01299; EPHRIN.1.
 CC Developmental protein; Neurogenesis; Glycoprotein; GPI-anchor; signal;
 CC Alternative splicing.
 CC FT SIGNAL 1 20 POTENTIAL.
 CC CHAIN 21 228 EPHRIN-A5.
 CC CARBOHYD 37 37 N-LINKED (GLCNAC...) (POTENTIAL).
 CC FT VARSLIC MISSING (IN SHORT ISOFORM).
 CC SQ SEQUENCE 228 AA; 26339 MW; 85439F5337420022 CRC64;
 CC -----
 CC Query Match 9.1%; Score 167.5; DB 1; Length 228;
 CC Best Local Similarity 28.8%; Pred. No. 0.00019;
 CC Matches 65; Conservative 29; Mismatches 81; Indels 51; Gaps 11;
 CC -----
 CC QY 33 VYVNSANKRFOAEAGYVLPQIGRLDILCPAPRPPGSHSSPNVEFYKLYLVG----- 85
 CC DB 34 VYVNSNSNPRFO-RGDYHIDVLCINDYLVCPHYEDSVPEDKT--ERYVLVYVNFQGSAC 90
 CC QY 86 -----GAQGRRCAPAPNLLTCORPDLILFTTKFOYSFNLWGHFRSHDYIYIAT 140
 CC DB 91 DHTSKGFKEWCKNRHPSN-----GPIKFSEKFOLETFPSLGFRRGRREYFYIS 141
 CC QY 141 S---DGTREGLESLQGVCLTRGMKYLRLVQGSPRGAVPRKPVSEMERDRGAASLE 197
 CC DB 142 AIPDNGRRS-----CLK--LKVFVR-----PTNSCKMTIGVDRVVDVNDKVENSL 186
 CC QY 198 PKENLPDPTSNATSRGAEGPLPPSPMPAVAGAAGLALLILGVA 243
 CC DB 187 PADDTV---HESAPSRG-BNNAQTPTIPSRL-----LAILFLLA 223
 CC -----
 CC RESULT 22
 CC EFNA5_RAT STANDARD; PRT; 228 AA.
 CC AC P97605;
 CC DT 01-NOV-1997 (Rel. 35; Created)
 CC DT 01-NOV-1997 (Rel. 35; Last sequence update)
 CC DT 30-MAY-2000 (Rel. 39; Last annotation update)
 CC DE Ephrin-A5 precursor (Eph-related receptor tyrosine kinase ligand 7)
 CC (LEK-7) (AL-1).
 CC GN EFNA5 OR EPLG7 OR LEK7 OR EPL7.
 CC OS Mus musculus (Mouse).

DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Ephrin-A5 precursor (EPH-related receptor tyrosine kinase ligand 7)
 DE (LRRK-7) (AL-1)
 GN EPNAS OR EPLG7 OR LERK7.
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=95267434; PubMed=7748564;
 RA Winkler J.W., Moran P., Valverde J., Shih A., Yuan J.Q., Wong S.C.,
 RA Tsai S.P., Goddard A., Henzel W.J., Hefti F., Beck K.D., Caras I.W.,
 RT "Cloning of AL-1, a ligand for an EPH-related tyrosine kinase
 RT receptor involved in axon bundle formation.";
 RL Neuron 14:973-981(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA Li Y.Y., McTiernan C.F., Feldman A.M.;
 RT LERK7, rat ligand for EPH-related receptor tyrosine kinase.";
 RL Submitted (Mar-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: MAY FUNCTION ACTIVELY TO STIMULATE AXON FASCICULATION.
 CC INDICES COMPARTMENTALIZED SIGNALING WITHIN A CAVOLAR-LIKE
 CC MEMBRANE MICRODOMAIN WHEN BOUND TO THE EXTRACELLULAR DOMAIN OF ITS
 CC COGNATE RECEPTOR. THIS SIGNALING EVENT REQUIRES THE ACTIVITY OF
 CC THE FYN TYROSINE KINASE (BY SIMILARITY).
 CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA2, EPHA3 AND
 CC EPHB1 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR IT
 CC IS COMPARTMENTALIZED IN DISCRETE CAVOLAR-LIKE MEMBRANE
 CC MICRODOMAINS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, HEART, PLACENTA AND LUNG.
 CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL: U69279; AAC05801.1; -
 DR InterPro: IPR001799; EPHrin.
 DR Pfam: PF00812; EPHrin.1.
 DR PRINTS: PR01347; EPHRIN.
 DR ProDom: PD002533; EPHrin.1.
 DR PROSITE: PS01299; EPHRIN.1.
 KW Developmental protein; Neurogenesis; Glycoprotein; GPI-anchor; Signal.
 FT SIGNAL 1 20
 FT CHAIN 1 228
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 228 AA; 26358 MW; 855985532D580022 CRC64;
 Query Match 9.1%; Score 167.5; DB 1; Length 228;
 Best Local Similarity 28.8%; Pred. No. 0.00019;
 Matches 65; Conservative 29; Mismatches 81; Indels 51; Gaps 11;
 QY 33 VYVNSANKRFOAEGGYVLPQIGRLDLICPPARPPGPHSSPNYERPKYLVLVG----- 85
 DB 34 VYVNSNSNPFQ-RGDYHIVCINDYLVDFCPHYEDSVPEDKT--EAYVLVYVNFQYSAC 90
 QY 86 -----GAQGRRCAPAPNILLTCRPPDLRTTIFQYISYPLNGHERSHDYIAT 140
 DB 91 DHTSKQFKRWECKRRPHSPN-----GPKRSEKFOULFTFSLGFERRPREYFYSS 141
 QY 141 S---DGTREGLESLGGVCLTRGKTYLVRVGGSPRGAVPRKPVSEMPERDRGAASILE 197
 DB 142 AITPDNGRRS-----CLK--LKYFVR-----PTMSCKKTIQVDRVYDVNDKXENSE 186

QY 198 PGKENTPGDPTNSATKCAEGPLPPSPMPAVAGAGLALLIGVA 243
 DB 187 PADDTV---HESAPSPRG-ENNAQTPIPERL-----LAILLFLLA 223
 RESULT 23
 EFAL HUMAN STANDARD; PRT; 205 AA.
 ID EFAL HUMAN
 AC P20827;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ephrin-A1 precursor (EPH-related receptor tyrosine kinase ligand 1)
 DE (LRRK-1) (Immediate-early response protein B61) (Tumor necrosis
 DE factor, alpha-induced protein 4).
 GN EPNAL OR EPLG1 OR LERK1 OR TNFALP4.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91042512; PubMed=2233719;
 RA Holzman L.B., Marks R.M., Dixie V.M.;
 RT "A novel immediate-early response gene of endothelium is induced by
 RT cytokines and encodes a secreted protein.";
 RL Mol. Cell. Biol. 10:5830-5838(1990).
 RN [2]
 RP GPI-ANCHOR.
 RX MEDLINE=95140419; PubMed=7838529;
 RA Kozlowski C.J., Maraskovsky E., McGrew J.T., Vandenbos T.,
 RA Teepe W., Lyman S.D., Srinivasan S., Fletcher P.A., Gayle R.B. III,
 RA Carrecci D.P., Beckmann M.P.;
 RT "Ligands for the receptor tyrosine kinases hek and elk: isolation of
 RT cDNAs encoding a family of proteins.";
 RL Oncogene 10:299-306(1995).
 CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA2, EPHA4,
 CC EPHA5, EPHA6 AND EPHA7. ALSO BINDS WITH LOW AFFINITY TO EPHA1.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- INDUCTION: BY TNF-ALPHA AND INTERLEUKIN-1 BETA.
 CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL: M57730; AA58388.1; -
 DR PIR: A36377; A36377.
 DR Genew; M8NC3221; EPNAL.
 DR MIM; 191164; -
 DR InterPro: IPR001799; EPHrin.
 DR Pfam: PF00812; EPHrin.1.
 DR PRINTS: PR01347; EPHRIN.
 DR ProDom: PD002533; EPHrin.1.
 DR PROSITE: PS01299; EPHRIN.1.
 KW Glycoprotein; GPI-anchor; Signal.
 FT SIGNAL 1 17
 FT CHAIN 1 205
 FT CARBOHYD 18 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 205 AA; 23771 MW; 4FEBC6BF4C1251A9 CRC64;
 Query Match 9.0%; Score 166; DB 1; Length 205;
 Best Local Similarity 27.5%; Pred. No. 0.00021;
 Matches 46; Conservative 31; Mismatches 74; Indels 16; Gaps 5;
 QY 18 LIGTGLVSLGLESPYVNSANKRFOAEGGYVLPQIGRLDLICPPARPPGPHSSPN-- 75
 DB 8 LIGLCCSLAADNHTVFWNSNPKFRNE-DYTIHVQINDYVDIICHYE---DHSVAADA 63

```

OY 76 YEFKLYVGGAGRCRCPAPNLLTCDPDI---DLRFITKEQVSPNLMGHEFRSH 132
Db 64 MEQYILYVHEEHEVOLCOPOSKDQVRWQCNRPSPAKHPEKLSERKFORFTFTLGEKFEKG 123
OY 133 HDYIITSDGTREGLESLOGVCLTRGMKYLIVGSGPAGVAPRK 119
Db 124 HSYIYSKPIHQHEDR-----CLRLKVTYVSGKITSPQAHVNPQE 163

RESULT 24
ID_EFAL_MOUSE STANDARD; PRT; 205 AA.
AC P9753;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Ephrin-A1 precursor (EPH-related receptor tyrosine kinase ligand 1)
DE (LERK-1) (Immediate early response protein B61).
GN EPFAL OR EPGL1 OR LERK1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mistral; TISSUE=Brain;
RX MEDLINE=95405853; PubMed=7675446;
RA Takahashi H., Ikeda T.;
RT "Molecular cloning and expression of rat and mouse B61 gene:
RT implications on organogenesis.";
RL Oncogene 11:879-883(1995).
CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA2, EPHA4,
CC EPHA5, EPHA6 AND EPHA7. ALSO BINDS WITH LOW AFFINITY TO EPHA1 (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D38056; BAA07242.1; -
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin; 1.
DR PRINTS; PR01347; EPHRIN.
DR PRODOM; PD002533; Ephrin; 1.
DR PROSITE; PS01299; EPHRIN; 1.
KW Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 17
FT CHAIN 18 205
FT CARBOHYD 26 26
FT SEQUENCE 205 AA; 23718 MW; CB6DAB3DB56A6EAD CRC64;
SQ
Query Match 8.8%; Score 162.5; DB 1; Length 205;
Best local similarity 25.7%; Pred. No. 0.00036;
Matches 49; Conservative 33; Mismatches 86; Indels 23; Gaps 6;
OY 18 LIGVLGVSGLSLEPVVNSANKRFQAGGYLVYQIGDRLLICPRAPPGPHSSPNYE 77
Db 8 LIGLCCSLAAVDRIHIVFVNSNPKFRRE-DYTVHVLQNDYLDITCPHYEDSV-ADAAVE 65
OY 78 FYKLYLVGGAGRCRCPAPNLLTCDPDI---DLRFITKEQVSPNLMGHEFRSH 134
Db 66 RYSLVWVHGEVVTCEPOSKDQVRWQCNRPSPAKHPEKLSERKFORFTFTLGEKFEKGS 125
OY 135 YIATSDGTREGLESLOGVCLTRGMKYLIVGSGPAGVAPRK 119
Db 124 HSYIYSKPIHQHEDR-----CLRLKVTYVSGKITSPQAHVNPQE 163

```

```

Db 126 YIYKPIYHOE-----TOCLKXTVNGKITSPHAHVNPQ-----EKLQADPEVQ 174
OY 190 --RGAAHSLSP 198
Db 175 VLHSGHSAAP 195

RESULT 25
ID_EFAL_MOUSE STANDARD; PRT; 205 AA.
AC P52793; P97331;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 35, Last annotation update)
DE Ephrin-A1 precursor (EPH-related receptor tyrosine kinase ligand 1)
DE (LERK-1) (Immediate early response protein B61).
GN EPFAL OR EPGL1 OR LERK1 OR EPL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR;
RX MEDLINE=95405853; PubMed=7675446;
RA Takahashi H., Ikeda T.;
RT "Molecular cloning and expression of rat and mouse B61 gene:
RT implications on organogenesis.";
RL Oncogene 11:879-883(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Morris J.C., Chiarletta A., Morris G.E., Giannotti J., Caruso A.,
RA Hammett D.J., Finnerty H., Turner K., Wood C.R.;
RA Submitted (May-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97060319; PubMed=8903354;
RA Flemmiken A.M., Gale N.W., Yancopoulos G.D., Wilkinson D.G.;
RT "Distinct and overlapping expression patterns of ligands for
RT Eph-related receptor tyrosine kinases during mouse embryogenesis.";
RL Dev. Biol. 179:382-401(1996).
CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA2, EPHA4,
CC EPHA5, EPHA6 AND EPHA7. ALSO BINDS WITH LOW AFFINITY TO EPHA1.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D38146; BAA07344.1; -
DR EMBL; U26188; AAB67563.1; -
DR EMBL; U90662; AAB50237.1; -
DR MGD; MGI:103236; Efnal.
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin; 1.
DR PRINTS; PR01347; EPHRIN.
DR PRODOM; PD002533; Ephrin; 1.
DR PROSITE; PS01299; EPHRIN; 1.
KW Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 17
FT CHAIN 18 205
FT CARBOHYD 26 26
FT CONFLICT 74 74
FT CONFLICT 79 79
FT CONFLICT 81 81
FT CONFLICT 91 91

```

```

FT CONFLICT 94 94 R -> Q (IN REF. 1)
FT CONFLICT 112 112 T -> S (IN REF. 1)
FT CONFLICT 115 115 I -> T (IN REF. 1)
FT CONFLICT 138 138 S -> T (IN REF. 1)
FT CONFLICT 154 154 N -> S (IN REF. 1)
FT CONFLICT 156 156 Q -> H (IN REF. 1)
FT CONFLICT 159 159 V -> A (IN REF. 1)
FT CONFLICT 181 181 Y -> H (IN REF. 1)
FT CONFLICT 204 204 S -> T (IN REF. 1)
SQ SEQUENCE 205 AA; 23802 MW; 5ABF3AE2091E868 CRC64;

Query Match
Best Local Similarity 26.1%; Score 161; DB 1; Length 205;
Matches 43; Conservative 33; Mismatches 77; Indels 12; Gaps 4;

QY 18 ILGLVGLVSGLSLEPYVNSANKRFOAGGYVLPQIGRLDLCPRAPPGHSPNTE 77
DB 8 LGLGCSLAADRHIVFMNNSNPKRFE-DYIVHVLNDYLDICPHYEDDSV-ADAAHE 65
QY 78 FYLLVYGAQGRCEAPPAVNLITCDRDL--DLRFTIKQEQYSPLMGHFRSHHD 134
DB 66 RYLLVYVHEHGEVYACQPOSKQVRMNCNRPSPAKHGPEKLSKQRFPTPLGKFEKHS 125
QY 135 YVLIATSDGTRREGLESQGVCLTRGAKYLRLVQSGPRGAVPRK 179
DB 126 YVYISKPIYHOE-----SQCLKLTVNGKITNHPQAHVDFOE 163

RESULT 26
EPA4_MOUSE STANDARD; PRT; 206 AA.
ID EPA4_MOUSE
AC 008542; 055218; (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ephrin-A4 precursor (EPH-related receptor tyrosine kinase ligand 4)
GN (LERK-4).
OS ERNA4 OR EPUG4 OR LERK4 OR EPL4.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=97060319; PubMed=8903354;
RX Fienmiken A.M., Gale N.W., Yancopoulos G.D., Wilkinson D.G.;
RA "Distinct and overlapping expression patterns of ligands for
RT Eph-related receptor tyrosine kinases during mouse embryogenesis.";
RL Dev. Biol. 179:382-401 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=98126446; PubMed=9465306;
RA Cerretti D.P., Nelson N.;
RT "Characterization of the genes for mouse LERK-3/Ephrin-A3 (Epl3),
RT mouse LERK-4/Ephrin-A4 (Epl4), and human LERK-6/Ephrin-A2 (Epl6);
RT conservation of intron/exon structure.";
RL Genomics 47:131-135 (1998).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U96663; AAB50238.1; -
CC EMBL: U92890; AAC39962.1; -
CC EMBL: U92889; AAC39962.1; JOINED.

```

```

DR MGI:106643; Efn4.
DR InterPro: IPR001799; Ephrin.
DR Pfam: PF00812; Ephrin.1.
DR PRINTS: PR01347; EPHRIN.
DR PRODOM: PD002533; Ephrin.1.
DR PROSITE: PS01299; EPHRIN.1.
KM Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 25
FT CHAIN 26 206
FT CARBOHYD 33 33 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SITE 41 43 CELL ATTACHMENT SITE (POTENTIAL).
FT CONFLICT 1 4
FT FT WRLL -> MLRLGLVPTPRPAPPPV (IN REF. 1).
SQ SEQUENCE 206 AA; 22861 MW; 43501971DDIC6EAS CRC64;

Query Match
Best Local Similarity 28.1%; Score 160.5; DB 1; Length 206;
Matches 61; Conservative 16; Mismatches 81; Indels 59; Gaps 10;

QY 32 PYVNSANKRFOAGGYVLPQIGRLDLCPRAPPGHSPNTEFYKLYVGAQGR 91
DB 29 PLYNNSNPRL-LRGDAVVELGFNDYLDICPHYESGPEGF--ETPALYMDWMSGYEA 85
QY 92 CEAPPAVNLITCDRDL--DLRFTIKQEQYSPLMGHFRSHHDYLIATSDGTRG 147
DB 86 CTAGANAFORWNCMPFAPSPVRFSEKIQRVTFPLGFEFLGRTYIIVPTPESG 145
QY 148 -LESLOGVCLTRGAKYLRLVQSGPRGAVPRKVSMPERDRAAH-SLEGKENTLP 205
DB 146 RCLMLQVSVCC-----KEGSSHEAHN-----VG 170
QY 206 DPTSNANS--RGAEGLPPSPMPAVNAGAGLALLL 240
DB 171 SPGSSGTSGMRGAPSP-----LCILL 194

RESULT 27
CAIA_MOUSE STANDARD; PRT; 680 AA.
ID CAIA_MOUSE
AC 005306;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Collagen alpha 1(X) chain precursor.
GN COL10A1.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=BALB/c;
RX MEDLINE=93143676; PubMed=8424763;
RA Elima K., Berola I., Rosati R., Metsaranta M., Garofalo S., Peralta M.,
RA de Crombrughe B., Vuorio E.;
RT "The mouse collagen X gene: complete nucleotide sequence, exon
RT structure and expression pattern.";
RL Biochem. J. 289:247-253 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV; TISSUE=Liver;
RX MEDLINE=9328750; PubMed=8477738;
RA Kong R.Y.C., Kwan K.M., Lau E.T., Thomas J.T., Boot-Handford R.P.,
RA Grant M.E., Cheah K.S.E.;
RT "Intron-exon structure, alternative use of promoter and expression of
RT the mouse collagen X gene, Col10a-1.";
RL Eur. J. Biochem. 213:99-111 (1993).
RN [3]
RP SEQUENCE OF 51-680 FROM N.A.
RC STRAIN=DAR/2J;
RX MEDLINE=92267014; PubMed=1587271;
RA Apke S.S., Seldin M.F., Hayaishi M., Olsen B.R.;

```


RP SEQUENCE OF 803-947.
 RX MEDLINE=80026030; PubMed=488910;
 RA Dews H., Fietzek P.P., Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. V. The amino
 RT acid sequence of the cyanogen bromide peptide alpha 1(III)C89A
 RT (position 789-927).";
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:851-860(1979).
 RN [61]
 RP SEQUENCE OF 948-1049.
 RX MEDLINE=80026031; PubMed=488911;
 RA Altmann H., Fietzek P.P., Glatville R.W., Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. VI. The amino
 RT acid sequence of the carboxyterminal cyanogen bromide peptide alpha
 RT 1(III)C89B (positions 928-1028).";
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:861-868(1979).
 RN [62]
 RP -1- FUNCTION: COLLAGEN TYPE III OCCURS IN MOST SOFT CONNECTIVE TISSUES
 CC ALONG WITH TYPE I COLLAGEN.
 CC -1- SUBUNIT: TRIMERS OF IDENTICAL ALPHA 1(III) CHAINS. THE CHAINS ARE
 CC LINKED TO EACH OTHER BY INTERCHAIN DISULFIDE BONDS. TRIMERS ARE
 CC ALSO CROSS-LINKED VIA HYDROXYLYSINES.
 CC -1- PPM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 DR PIR; A02862; C8B07S.
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF01391; Collagen_17.
 DR ProDom; PD000007; Collagen_1.
 DR PROSITE; PS01208; WVC; PARTIAL.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Glycoprotein; Collagen.
 FT DOMAIN 1 14 NONHELICAL REGION (N-TERMINAL).
 FT DOMAIN 15 1040 TRIPLE-HELICAL REGION.
 FT MOD_RES 1041 1049 NONHELICAL REGION (C-TERMINAL).
 FT MOD_RES 95 95 HYDROXYLATION.
 FT MOD_RES 107 107 HYDROXYLATION.
 FT MOD_RES 119 119 HYDROXYLATION.
 FT MOD_RES 938 938 HYDROXYLATION.
 FT MOD_RES 950 950 HYDROXYLATION.
 FT CARBOHYD 107 107 O-LINKED (GAL. . .).
 FT CARBOHYD 950 950 O-LINKED (GAL. . .).
 FT DISULFID 1040 1040 INTERCHAIN.
 FT DISULFID 1041 1041 INTERCHAIN.
 SQ SEQUENCE 1049 AA; 93651 MW; 88EC33D1C6BC9A3 CRC64;
 Query Match 8.4%; Score 154.5; DB 1; Length 1049;
 Best Local Similarity 26.5%; Pred. No. 0.00674; Indels 83; Gaps 8;
 Matches 60; Conservative 9; Mismatches 74;
 QY 171 PRGAVPRKVESEMPERDRGA-----ASHLEFGKEN 202
 DB 688 PAGESPAGSPGQGVKSGRSGPGAGFPGGRRPDPGSGNGNPGPSSGAPKDG 747
 QY 203 LRPDPISNAT-----SRGAEPLPPSPMPAVAGAAAGLALLIGVAGA 245
 DB 748 PGPSPSGNAPGSPGISGPKGDSGPPERGAPGPQGPFGAPGLGAG-----LTGARL 802
 QY 246 GGMWRRRRARRAK-----SESRHPP-----GSFGRGSGYLGGGGMGPREAPGE 292
 DB 803 AGPSPGMPGARGSPGQIKENKGRPSGNGRGPFGQGLPLAGTAGEPRGRDNPGS 862
 QY 293 LGIALRG-----GAADPPFCPHYKVSQDYGPVYIVODGPP 330
 DB 863 DGLPGRDGAAPGAKGRGNGSPGAP-----GAPGHP-----GPP 896
 RESULT 29
 CC02_CABEL STANDARD; RT; 301 AA.
 ID CC02_CABEL
 AC 217656;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cuticle collagen 2 precursor.

GN COL-2 OR W01B6.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodidae; Caenorhabditis.
 CX NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=83050944; PubMed=7139711;
 RA Krimer J.M., Cox G.N., Hirsch D.;
 RT "Comparisons of the complete sequences of two collagen genes from
 RT C. elegans.";
 RL Cell 30:559-606(1982).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=85105075; PubMed=2578467;
 RA Krimer J.M., Cox G.N., Hirsch D.;
 RT "Expression of the Caenorhabditis elegans collagen genes col-1 and
 RT col-2 is developmentally regulated.";
 RL J. Biol. Chem. 260:1945-1951(1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Coles L.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: NEMATODE CUTICLES ARE COMPOSED LARGELY OF COLLAGEN-LIKE
 CC PROTEINS. THE CUTICLE FUNCTIONS BOTH AS AN EXOSKELETON AND AS A
 CC BARRIER TO PROTECT THE NEM FROM ITS ENVIRONMENT.
 CC -1- SUBUNIT: COLLAGEN POLYPEPTIDE CHAINS ARE COMPLEXED WITHIN THE
 CC CUTICLE BY DISULFIDE BONDS AND OTHER TYPES OF COVALENT CROSS-
 CC LINKS.
 CC -1- SIMILARITY: BELONGS TO THE CUTICULAR COLLAGEN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; J01048; AAA27990.1; -
 DR EMBL; V00148; CAA23464.1; -
 DR EMBL; Z68301; CAA92620.1; -
 DR PIR; B31219; B31219
 DR WormPep; W01B6.7; CE03759.
 DR InterPro; IPR002486; Col_cuticle_N.
 DR Pfam; PF01391; Collagen_2.
 DR Pfam; PF01484; Col_cuticle_N_1.
 KW Cuticle; Connective tissue; Repeat; Multigene family; Collagen;
 FT SIGNAL 1 37 POTENTIAL.
 FT CHAIN 38 301 CUTICLE COLLAGEN 2.
 FT DOMAIN 85 97 GLY-RICH.
 FT DOMAIN 105 134 TRIPLE-HELICAL REGION.
 FT DOMAIN 153 176 TRIPLE-HELICAL REGION.
 FT DOMAIN 183 212 TRIPLE-HELICAL REGION.
 FT DOMAIN 215 282 TRIPLE-HELICAL REGION.
 SQ SEQUENCE 301 AA; 28025 MW; 33317E3BDAC302F9 CRC64;
 Query Match 8.2%; Score 151.5; DB 1; Length 301;
 Best Local Similarity 28.7%; Pred. No. 0.0027;
 Matches 50; Conservative 15; Mismatches 72; Indels 37; Gaps 7;
 QY 168 GSPRGATVPAKPVSEMPERDRG-----AAHSLPFGKENVPGPT--SNATSGAE 217
 DB 129 GNPBGASAPCEPTQPCPCPGPPGAPGPPGPDNDNGSPGAPSGPAGPP 188
 QY 218 GLLPSPMPAVAGAAAGLALLIGVAGA-----GGAMCWRRRARRAPSESRHGP--GSFGRG 273
 DB 189 GPAPGAGNDGAPGAPG-----PGEFASSEGGP-----GBPGAGPPGPPAGNDG 236

CY 274 SLGGGGGGMPPR-----EAEPELGIALRGGAADPPCFCHYKVSQ 316
 ID CA34 HUMAN STANDARD; PRT; 1670 AA.
 AC 001955; 09B0R2;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUN-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Collagen alpha 3 (IV) chain precursor (Goodpasture antigen).
 GN COL4A3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=94364994; PubMed=8083201;
 RA Matyama M., Leinonen A., Mochizuki T., Trygvasen K., Redders S.T.;
 RT "Complete primary structure of the human alpha 3(IV) collagen chain.
 RT Coexpression of the alpha 3 (IV) and alpha 4 (IV) collagen chains in
 RT human tissues.";
 RL J. Biol. Chem. 269:123013-23017 (1994).
 RN [2]
 RP REVISIONS.
 RA Leinonen A.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A., VARIANTS AS E-297; R-407; R-640; R-1167; E-1207;
 RP Q-1215; S-1277; T-1330; E-1334; E-1347 AND C-1661, AND VARIANTS R-43;
 RP E-162; Y-326; H-408; R-451; L-574; E-1269 AND P-1474.
 RX MEDLINE=21064666; PubMed=11134255;
 RA Heider L., Aronow C., Forestier L., Cohen-Solal L., Mollet G.,
 RA Guetierrez B., Savrou C., Gubler M.C., Antignac C.;
 RT "Structure of the human type IV collagen gene COL4A3 and mutations in
 RT autosomal Alport syndrome.";
 RL J. Am. Soc. Nephrol. 12:97-106 (2001).
 RN [4]
 RP SEQUENCE OF 1386-1670 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=93015826; PubMed=1400291.
 RA Quinones S., Bernal D., Garcia-Sogo M., Elena S.F., Saus J.;
 RT "Exon/intron structure of the human alpha 3(IV) gene encompassing the
 RT Goodpasture antigen (alpha 3(IV)NC1). Identification of a potentially
 RT antigenic region at the triple helix/NC1 domain junction.";
 RL J. Biol. Chem. 267:19780-19784 (1992).
 RN [5]
 RP SEQUENCE OF 1453-1670 FROM N.A.
 RX MEDLINE=91353701; PubMed=1882840.
 RA Morrison K.E., Matyama M., Yang-Feng T.L., Redders S.T.;
 RT "Sequence and localization of a partial cDNA encoding the human alpha
 RT 3 chain of type IV collagen.";
 RL Am. J. Hum. Genet. 49:545-554 (1991).
 RN [6]
 RP SEQUENCE OF 1331-1670 FROM N.A.
 RX TISSUE=Kidney;
 RX MEDLINE=92147878; PubMed=1737849;
 RA Turner N., Mason P.J., Brown R., Fox M., Povey S., Rees A.,
 RA Pusey C.D.;
 RT "Molecular cloning of the human Goodpasture antigen demonstrates it
 RT to be the alpha 3 chain of type IV collagen.";
 RL J. Clin. Invest. 89:592-601 (1992).
 RN [7]
 RP SEQUENCE OF 1644-1670 FROM N.A.
 RX TISSUE=Kidney;
 RX Ding J.;
 RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE OF 1439-1670, AND ALTERNATIVE SPLICING.

RC TISSUE=Kidney;
 RX MEDLINE=94124597; PubMed=8294492;
 RA Feng L., Xia Y., Wilson C.B.;
 RT "Alternative splicing of the NC1 domain of the human alpha 3(IV)
 RT collagen gene. Differential expression of mRNA transcripts that
 RT predict three protein variants with distinct carboxyl regions.";
 RL J. Biol. Chem. 269:2342-2348 (1994).
 RN [9]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE=96196854; PubMed=9537506;
 RA Momota R., Sugimoto M., Ohashi T., Kigasawa K., Yoshioke H.,
 RA Ninomiya Y.;
 RT "Two genes, COL4A3 and COL4A4 coding for the human alpha3 (IV) and
 RT alpha4 (IV) collagen chains are arranged head-to-head on chromosome
 RT 2q36.";
 RL FEBS Lett. 424:11-16 (1998).
 RN [10]
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=93280184; PubMed=8505332;
 RA Bernal D., Quinones S., Saus J.;
 RT "The human mRNA encoding the Goodpasture antigen is alternatively
 RT spliced.";
 RL J. Biol. Chem. 268:12090-12094 (1993).
 RN [11]
 RP VARIANT PRO-1474.
 RX MEDLINE=95078827; PubMed=7987301;
 RA Lemmink H.H., Mochizuki T., van den Heuvel L.P.M.J., Schroeder C.H.,
 RA Barrientos A., Monens L.A.H., van Oost B.A., Brunner H.G.,
 RA Redders S.T., Smeets H.J.M.;
 RT "Mutations in the type IV collagen alpha 3 (COL4A3) gene in autosomal
 RT recessive Alport syndrome.";
 RL Hum. Mol. Genet. 3:1269-1273 (1994).
 CC -1- FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF
 CC GLOMERULAR BASEMENT MEMBRANES (GBM), FORMING A "CHICKEN-WIRE"
 CC MESHWORK TOGETHER WITH LAMININS, PROTEGLYCANS AND ENACTIN/
 CC NIDAGEN.
 CC -1- SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV) -
 CC ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE
 CC WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
 CC -1- SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL).
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2/V AND
 CC 3/L5; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER IN THEIR
 CC C-TERMINAL NC1 DOMAINS.
 CC -1- TISSUE SPECIFICITY: ALPHA 3 AND ALPHA 4 TYPE IV COLLAGENS ARE
 CC COLocalized AND PRESENT ONLY IN BASEMENT MEMBRANES OF KIDNEY, EYE,
 CC COCHLEA, LUNG AND BRAIN.
 CC -1- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
 CC DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
 CC G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
 CC CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
 CC TRIPLE-HELICAL 7S DOMAIN.
 CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPETIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -1- PTM: THE ALTERNATIVE SPLICED FORM V CONTAINS AN ADDITIONAL
 CC N-LINKED GLYCOSYLATION SITE.
 CC -1- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH
 CC ARE INVOLVED IN INTER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF
 CC THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE
 CC IV COLLAGENS.
 CC -1- PTM: Phosphorylated by the Goodpasture antigen-binding protein.
 CC -1- DISEASE: ANTIBODIES AGAINST THE NC1 DOMAIN OF ALPHA3(IV) MEDIATE
 CC THE HUMAN AUTOIMMUNE DISEASE, GOODPASTURE SYNDROME, WHICH IS
 CC CHARACTERIZED BY HEMATURIA AND PULMONARY HEMORRHAGE.
 CC -1- DISEASE: DEFECTS IN COL4A3 ARE ASSOCIATED WITH THE TYPE I
 CC AUTOSOMAL RECESSIVE FORM OF ALPORT SYNDROME, AN HEREDITARY
 CC GLOMERULONEPHROPATHY CHARACTERIZED BY PROGRESSIVE RENAL FAILURE,
 CC HEMATURIA AND DEAFNESS. THE RECESSIVE FORM OCCURS EQUALLY BETWEEN
 CC MALES AND FEMALES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

CC
 DR EMBL; X60031; CAA56335.1; -
 DR EMBL; AJ288487; CAC36101.1; JOINED.
 DR EMBL; AJ288488; CAC36101.1; JOINED.
 DR EMBL; AJ288489; CAC36101.1; JOINED.
 DR EMBL; AJ288490; CAC36101.1; JOINED.
 DR EMBL; AJ288491; CAC36101.1; JOINED.
 DR EMBL; AJ288492; CAC36101.1; JOINED.
 DR EMBL; AJ288493; CAC36101.1; JOINED.
 DR EMBL; AJ288494; CAC36101.1; JOINED.
 DR EMBL; AJ288495; CAC36101.1; JOINED.
 DR EMBL; AJ288496; CAC36101.1; JOINED.
 DR EMBL; AJ288497; CAC36101.1; JOINED.
 DR EMBL; AJ288498; CAC36101.1; JOINED.
 DR EMBL; AJ288499; CAC36101.1; JOINED.
 DR EMBL; AJ288500; CAC36101.1; JOINED.
 DR EMBL; AJ288501; CAC36101.1; JOINED.
 DR EMBL; AJ288502; CAC36101.1; JOINED.
 DR EMBL; AJ288503; CAC36101.1; JOINED.
 DR EMBL; AJ288504; CAC36101.1; JOINED.
 DR EMBL; AJ288505; CAC36101.1; JOINED.
 DR EMBL; AJ288506; CAC36101.1; JOINED.
 DR EMBL; AJ288507; CAC36101.1; JOINED.
 DR EMBL; AJ288508; CAC36101.1; JOINED.
 DR EMBL; AJ288509; CAC36101.1; JOINED.
 DR EMBL; AJ288510; CAC36101.1; JOINED.
 DR EMBL; AJ288511; CAC36101.1; JOINED.
 DR EMBL; AJ288512; CAC36101.1; JOINED.
 DR EMBL; AJ288513; CAC36101.1; JOINED.
 DR EMBL; AJ288514; CAC36101.1; JOINED.
 DR EMBL; AJ288515; CAC36101.1; JOINED.
 DR EMBL; AJ288516; CAC36101.1; JOINED.
 DR EMBL; AJ288517; CAC36101.1; JOINED.
 DR EMBL; AJ288518; CAC36101.1; JOINED.
 DR EMBL; AJ288519; CAC36101.1; JOINED.
 DR EMBL; AJ288520; CAC36101.1; JOINED.
 DR EMBL; AJ288521; CAC36101.1; JOINED.
 DR EMBL; AJ288522; CAC36101.1; JOINED.
 DR EMBL; AJ288523; CAC36101.1; JOINED.
 DR EMBL; AJ288524; CAC36101.1; JOINED.
 DR EMBL; AJ288525; CAC36101.1; JOINED.
 DR EMBL; AJ288526; CAC36101.1; JOINED.
 DR EMBL; AJ288527; CAC36101.1; JOINED.
 DR EMBL; AJ288528; CAC36101.1; JOINED.
 DR EMBL; AJ288529; CAC36101.1; JOINED.
 DR EMBL; AJ288530; CAC36101.1; JOINED.
 DR EMBL; AJ288531; CAC36101.1; JOINED.
 DR EMBL; AJ288532; CAC36101.1; JOINED.
 DR EMBL; AJ288533; CAC36101.1; JOINED.
 DR EMBL; AJ288534; CAC36101.1; JOINED.
 DR EMBL; AJ288535; CAC36101.1; JOINED.
 DR EMBL; AJ288536; CAC36101.1; JOINED.
 DR EMBL; AJ288537; CAC36101.1; JOINED.
 DR EMBL; AJ288538; CAC36101.1; JOINED.
 DR EMBL; M92993; AAA21610.1; -
 DR EMBL; S55790; AAB19637.1; -
 DR EMBL; M81379; AAA51556.1; -
 DR EMBL; L08650; AAA52044.1; -
 DR EMBL; U02519; AAA18942.1; -

Query Match 8.0%; Score 148.5; DB 1; Length 1670;
 Best Local Similarity 27.1%; Pred. No. 0.023;
 Matches 105; Conservative 22; Mismatches 146; Indels 115; Gaps 23;

OY 1 MGPFRHS-GPGGV-----RVGALTLGLVGLVSGLSLEPVYWNANRKRFOLEGGYLVLPQI 54
 DB 1020 MGNMGSKSKKCTGLGFRPGRAGPGI-PGIHGL-QGDKXEPGV-----SEG----- 1061
 OY 55 GDRIDLICPRAPRPPGPHSSPNYEPYKLYLVGAQGRRC-----APPA 97

DB 1062 -----TRGPPGPTGDPG-----LPGDMGKXGEMQPPGPIGLPAGPEGAGS 1105
 OY 98 PNL-----LILCDRPPDLIRPT-IKFOEYSPNLMGHE-----FRSHHYIYIATSDGREGLES 150
 DB 1106 PPSPELPEKPPGPHDLDGKGLIKGLIGPPGIRGPPGLPFPSPGPPGWRGDQG-RDGI PG 1164
 OY 151 LOGGVCLTRGKVVLLRYQSPRG--GAVPRKPVSEMMERDRGAHSLPEKENVPGDPT 208
 DB 1165 PAG-----EKGETGLRAPRPPGRNPGAQAK-----GDRGA-----PGPGLPG--- 1204
 OY 209 SNATSRGAEGLPPSPMAYVAGAGLALLLVAGAGMCMRRRAKPSBSHRPG-- 266
 DB 1205 -RKGMAGAGRPGPTGIRGPPGPPGLPGAILPGGTGRGP--PGSRGSPAPRPPGPG 1260
 OY 267 -----GSPFRGSGLG-LGGGGGMP--RAEPGELGIALRGGAADPPPCPHYEK 313
 DB 1261 SHVIGIKGDKSMHGRPKGPGGTAGDMGPRGRIGAROTGLP---GPRGDPG-QGFR PG 1316
 OY 314 VSGDYGHFVYIVQDP-----PQSPENI 336
 DB 1317 VKGEKGNPGFLGSGPPGPIGPKGPPGV 1344

RESULT 31
 ID CAH_MOUSE STANDARD, PRT; 1527 AA.
 AC P39061; Q62002; Q61437;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DS Collagen alpha 1(XVIII) chain precursor [contains: Endostatin].
 GN COL18A1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).
 RC STRAIN=BALE/c; TISSUE=Liver;
 RX MEDLINE=94245707; PubMed=8186673;
 RA Rehn M.V., Hintikka E., Pihlajaniemi T.;
 RT "Primary structure of the alpha 1 chain of mouse type XVIII collagen,
 RT partial structure of the corresponding gene, and comparison of the
 RT alpha 1(XVIII) chain with its homologue, the alpha 1(XV) collagen
 RT chain.";
 RT J. Biol. Chem. 269:13929-13935(1994).
 RN [2]
 RP SEQUENCE FROM N.A. (SHORT AND LONG ISOFORMS).
 RA Rehn M., Hintikka E., Pihlajaniemi T.;
 RT "Characterization of the mouse gene for the alpha-1 chain of type
 RT XVIII collagen (COL18A1) reveals that the three variant N-terminal
 RT polypeptide forms are transcribed from two widely separated
 RT promoters.";
 RT Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 213-1140 FROM N.A. (SHORT ISOFORM).
 RX MEDLINE=94240112; PubMed=8183894;
 RA Rehn M.V., Pihlajaniemi T.;
 RT "Alpha 1(XVIII), a collagen chain with frequent interruptions in the
 RT collagenous sequence, a distinct tissue distribution, and homology
 RT with type XV collagen.";
 RT Proc. Natl. Acad. Sci. U.S.A. 91:4234-4238(1994).
 RN [4]
 RP SEQUENCE OF 240-1527 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=94240111; PubMed=8183893;
 RA Oh S.P., Kamagata Y., Muzigaki Y., Timmons S., Ooshima A., Olsen B.R.;
 RT "Isolation and sequencing of cDNAs for proteins with multiple domains
 RT of Gly-Xaa-Yaa repeats identify a distinct family of collagenous
 RT proteins.";
 RT Proc. Natl. Acad. Sci. U.S.A. 91:4229-4233(1994).
 RN [5]
 RP CHARACTERIZATION OF ENDOSTATIN, AND PARTIAL SEQUENCE.

EX MEDLINE=97160848; PubMed=9008168;
 RA O'Reilly M.S., Boehm T., Shing Y., Fukai N., Vasios G., Lane W.S.,
 RA Flynn E., Birkhead J.R., Olsen B.R., Folkman J.;
 RT "Endostatin: an endogenous inhibitor of angiogenesis and tumor
 RT growth.";
 RL Cell 88:277-285(1997).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF ENDOSTATIN.
 RA MEDLINE=98169382; PubMed=9501087;
 RA Honenester E., Sasaki T., Olsen B.R., Timpl R.;
 RT "Crystal structure of the angiogenesis inhibitor endostatin at 1.5-A
 RT resolution.";
 RL EMBO J. 17:1656-1664(1998).
 CC -|- FUNCTION: ENDOSTATIN POTENTIALLY INHIBITS ENDOTHELIAL CELL
 CC PROLIFERATION AND ANGIOGENESIS. MAY INHIBIT ANGIOGENESIS BY
 CC BINDING TO THE HEPARAN SULPHATE PROTEOGLYCANS INVOLVED IN GROWTH
 CC FACTOR SIGNALING.
 CC -|- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SHORT FORM AND A LONG FORM
 CC (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -|- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPETIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -|- SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH
 CC INTERRUPTED HELICES (FACIT) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L16898; AAA37434.1; -;
 DR EMBL; U03714; AAA20657.1; -;
 DR EMBL; U03715; AAC52901.1; -;
 DR EMBL; U34606; AAC52901.1; JOINED.
 DR EMBL; U34608; AAC52901.1; JOINED.
 DR EMBL; U34609; AAC52901.1; JOINED.
 DR EMBL; U34610; AAC52901.1; JOINED.
 DR EMBL; U34611; AAC52901.1; JOINED.
 DR EMBL; U34612; AAC52901.1; JOINED.
 DR EMBL; U03716; AAC52901.1; JOINED.
 DR EMBL; U03718; AAC52901.1; JOINED.
 DR EMBL; U03715; AAC52902.1; -;
 DR EMBL; U34607; AAC52902.1; JOINED.
 DR EMBL; U34608; AAC52902.1; JOINED.
 DR EMBL; U34609; AAC52902.1; JOINED.
 DR EMBL; U34610; AAC52902.1; JOINED.
 DR EMBL; U34611; AAC52902.1; JOINED.
 DR EMBL; U34612; AAC52902.1; JOINED.
 DR EMBL; U03716; AAC52902.1; JOINED.
 DR EMBL; U03718; AAC52902.1; JOINED.
 DR EMBL; U11636; AAC52902.1; -;
 DR EMBL; U12545; AAA19787.1; -;
 DR PDB; 1XOE; 16-FEB-99.
 DR MGD; MGI:88451; Coll18a1.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR001791; Laminin_G.
 DR InterPro; IPR003129; TSPN.
 DR Pfam; PF01391; Collagen; 8.
 DR Pfam; PF02210; TSPN; 1.
 DR SMART; SM00282; LamG; 1.
 DR SMART; SM00282; LamG; 1.
 DR SMART; SM00210; TSPN; 1.
 DR Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Cell adhesion; Collagen; Glycoprotein; Signal; Alternative splicing;
 KW 3D-structure.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 1527 COLLAGEN ALPHA 1(XVIII) CHAIN.
 FT CHAIN 1344 1527 ENDOSTATIN.
 FT DOMAIN 27 538 NONHEMICAL REGION 1 (NC1).
 FT DOMAIN 539 565 TRIPLE-HEMICAL REGION 1 (COL1).

FT	DOMAIN	566	575	NONHEMICAL REGION 2 (NC2).
FT	DOMAIN	576	649	TRIPLE-HEMICAL REGION 2 (COL2).
FT	DOMAIN	650	673	NONHEMICAL REGION 3 (NC3).
FT	DOMAIN	674	795	TRIPLE-HEMICAL REGION 3 (COL3).
FT	DOMAIN	796	818	NONHEMICAL REGION 4 (NC4).
FT	DOMAIN	819	901	TRIPLE-HEMICAL REGION 4 (COL4).
FT	DOMAIN	902	915	NONHEMICAL REGION 5 (NC5).
FT	DOMAIN	916	957	TRIPLE-HEMICAL REGION 5 (COL5).
FT	DOMAIN	958	970	NONHEMICAL REGION 6 (NC6).
FT	DOMAIN	971	1043	TRIPLE-HEMICAL REGION 6 (COL6).
FT	DOMAIN	1044	1053	NONHEMICAL REGION 7 (NC7).
FT	DOMAIN	1054	1086	TRIPLE-HEMICAL REGION 7 (COL7).
FT	DOMAIN	1087	1098	NONHEMICAL REGION 8 (NC8).
FT	DOMAIN	1099	1122	TRIPLE-HEMICAL REGION 8 (COL8).
FT	DOMAIN	1123	1129	NONHEMICAL REGION 9 (NC9).
FT	DOMAIN	1130	1181	TRIPLE-HEMICAL REGION 9 (COL9).
FT	DOMAIN	1182	1194	NONHEMICAL REGION 10 (NC10).
FT	DOMAIN	1195	1212	TRIPLE-HEMICAL REGION 10 (COL10).
FT	DOMAIN	1213	1527	NONHEMICAL REGION 11 (NC11).
FT	CARBOHYD	338	338	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	700	700	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	1376	1516	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	1478	1508	
FT	SITE	1104	1106	
FT	VARSPLIC	213	212	CELL ATTACHMENT SITE (POTENTIAL).
FT	VARSPPLIC	213	212	MISSING (IN SHORT ISOFORM).
FT	CONFLICT	900	900	APPTOLPFGOSNLQAPRPAPDPF -> MAPRWHLIDVL
FT	CONFLICT	947	947	TSVLIVARVSMAE (IN SHORT ISOFORM).
FT	CONFLICT	964	964	P -> F (IN REF. 4).
FT	CONFLICT	1157	1157	A -> R (IN REF. 4).
FT	CONFLICT	1266	1266	R -> P (IN REF. 4).
FT	CONFLICT	1276	1276	P -> L (IN REF. 4).
FT	CONFLICT	1437	1437	L -> F (IN REF. 4).
FT	CONFLICT	1527	1527	L -> V (IN REF. 4).
SO	SEQUENCE	1527 AA;	156008 MM;	964504SAFL140B513 CRC64;

Query Match 7.9%; Score 146.5; DB 1; Length 1527;
 Best Local Similarity 23.7%; Pred. No. 0.029;
 Matches 92; Conservative 22; Mismatches 107; Indels 167; Gaps 21;

QY	2	GPHSGPGGVRYVALLLVGLG-LVSGLSLSEPVYVNMNSANKRQAGGYVLPIQIDRLDL	60
DB	551	GPP--GPAGQ-----GPAGVVOGSPNSQPV-----	574
QY	61	LCPRAR--PGPHSSPNYEFYKLYLVGAQGRCE-----	93
DB	575	--PGAGGPPGPPGPKD-----GTRGRDGPEDGPDGDRPGDTGQGPPTGPDVG	624
QY	94	-----APPAPNLLTCDRDLDFRTTKFQSYSPNLMGHERSHHDYIAT	140
DB	625	PKGEKDPGIGRGPFGP-----GPPGPPSPRQDKLTFTDME	661
QY	141	SDTRGSLGSLGGLVTRGMKYLRLVQSPRGAAYPRKP--VSEMPMERDR-GAAHSLE	197
DB	662	GSFGSDIESLRG-----PRGPPGPPPPPPVPLGPPGPPGFGINGSYA	704
QY	198	PGKENPDPPTSNATSGAEGP--LP-PPSPAPVAGAACGLALLLGVAGAGCANCRRR	254
DB	705	PGPAGPPGVP-----GREGPPGPPGPPGPPGPKKEP-----FGVAGQGSV	747
QY	255	RAPSSRRHGP--GSRGRGSLGLGGGGMGRBAPGSLGIALGG--GAADPPPCPHY	311
DB	748	----GVGIFPGKSGKDGIPGMPKSGIA--GSFPGVPGPPGPPGPPGPPGFAAGF	800
QY	312	EKVSQDYGHVYIVQ-----DGPQSP	333
DB	801	DDWEGS-GIPLWTTARSDGLGPPGSP	827

RESULT 32
 CAFF_RIFPA STANDARD; PRT; 1027 AA.
 ID CAFF_RIFPA
 AC P30754;

DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Fibril-forming collagen alpha chain.
 OS Riffia pachypoda (tube worm).
 OC Eukaryota; Metazoa; Pogonophora; Vestimentifera; Axonobranchia;
 OC Riffidae; Riffidae; Riffia.
 OK NCBI_TaxId=6426;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=93130909; PubMed=1483468;
 RA Mann K., Gaill F., Timpl R.;
 RT "Amino-acid sequence and cell-adhesion activity of a fibril-forming
 RT collagen from the tube worm Riffia pachypoda living at deep sea
 RT hydrothermal vents.";
 RL Eur. J. Biochem. 210:839-847 (1992).
 RN [2]
 RP SEQUENCE OF 8-45; 525-618 AND 810-882.
 RC TISSUE-Cuticle;
 RX MEDLINE=92015309; PubMed=1920405;
 RA Gaill F., Wiedemann H., Mann K., Kuhn K., Timpl R., Engel U.;
 RT "Molecular characterization of cuticle and interstitial collagens
 RT from worms collected at deep sea hydrothermal vents.";
 RL J. Mol. Biol. 221:209-223 (1991).
 CC -1- FUNCTION: FIBRIL-FORMING COLLAGEN.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- PTM: ALL THE "X" ARE PROBABLY GLYCOSYLATED HYDROXYLISINES.
 DR PIR; S22915; S22915.
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF01391; Collagen; 16.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Glycoprotein; Collagen.
 FT DOMAIN 1 12
 FT DOMAIN 13 1023 NONHELICAL REGION (N-TERMINAL).
 FT DOMAIN 1024 1027 TRIPLE-HELICAL REGION.
 FT MOD RES 21 21 NONHELICAL REGION (C-TERMINAL).
 FT MOD RES 24 24 HYDROXYLATION (PARTIAL).
 FT MOD RES 27 27 HYDROXYLATION (PARTIAL).
 FT MOD RES 39 39 HYDROXYLATION.
 FT MOD RES 53 53 HYDROXYLATION.
 FT MOD RES 54 54 HYDROXYLATION (PARTIAL).
 FT MOD RES 72 72 HYDROXYLATION (PARTIAL).
 FT MOD RES 90 90 HYDROXYLATION.
 FT MOD RES 93 93 HYDROXYLATION.
 FT MOD RES 123 123 HYDROXYLATION (PARTIAL).
 FT MOD RES 128 128 HYDROXYLATION (PARTIAL).
 FT MOD RES 150 150 HYDROXYLATION.
 FT MOD RES 161 161 HYDROXYLATION (PARTIAL).
 FT MOD RES 162 162 HYDROXYLATION.
 FT MOD RES 164 164 HYDROXYLATION (PARTIAL).
 FT MOD RES 165 165 HYDROXYLATION.
 FT MOD RES 174 174 HYDROXYLATION.
 FT MOD RES 177 177 HYDROXYLATION.
 FT MOD RES 180 180 HYDROXYLATION.
 FT MOD RES 183 183 HYDROXYLATION.
 FT MOD RES 207 207 HYDROXYLATION.
 FT MOD RES 216 216 HYDROXYLATION.
 FT MOD RES 219 219 HYDROXYLATION.
 FT MOD RES 228 228 HYDROXYLATION.
 FT MOD RES 237 237 HYDROXYLATION.
 FT MOD RES 243 243 HYDROXYLATION (PARTIAL).
 FT MOD RES 249 249 HYDROXYLATION.
 FT MOD RES 255 255 HYDROXYLATION.
 FT MOD RES 273 273 HYDROXYLATION (PARTIAL).
 FT MOD RES 276 276 HYDROXYLATION (PARTIAL).
 FT MOD RES 285 285 HYDROXYLATION (PARTIAL).
 FT MOD RES 291 291 HYDROXYLATION (PARTIAL).
 FT MOD RES 303 303 HYDROXYLATION (PARTIAL).
 FT MOD RES 306 306 HYDROXYLATION.
 FT MOD RES 312 312 HYDROXYLATION.
 FT MOD RES 321 321 HYDROXYLATION.
 FT MOD RES 327 327 HYDROXYLATION.
 FT MOD RES 339 339 HYDROXYLATION.

FT MOD RES 342 342 HYDROXYLATION.
 FT MOD RES 348 348 HYDROXYLATION (PARTIAL).
 FT MOD RES 351 351 HYDROXYLATION (PARTIAL).
 FT MOD RES 366 366 HYDROXYLATION.
 FT MOD RES 372 372 HYDROXYLATION.
 FT MOD RES 375 375 HYDROXYLATION.
 FT MOD RES 381 381 HYDROXYLATION (PARTIAL).
 FT MOD RES 387 387 HYDROXYLATION.
 FT MOD RES 416 416 HYDROXYLATION (PARTIAL).
 FT MOD RES 417 417 HYDROXYLATION.
 FT MOD RES 423 423 HYDROXYLATION.
 FT MOD RES 429 429 HYDROXYLATION.
 FT MOD RES 432 432 HYDROXYLATION.
 FT MOD RES 433 433 HYDROXYLATION.
 FT MOD RES 465 465 HYDROXYLATION.
 FT MOD RES 483 483 HYDROXYLATION.
 FT MOD RES 500 500 HYDROXYLATION (PARTIAL).
 FT MOD RES 503 503 HYDROXYLATION (PARTIAL).
 FT MOD RES 506 506 HYDROXYLATION (PARTIAL).
 FT MOD RES 513 513 HYDROXYLATION.
 FT MOD RES 525 525 HYDROXYLATION.
 FT MOD RES 533 533 HYDROXYLATION (PARTIAL).
 FT MOD RES 536 536 HYDROXYLATION (PARTIAL).
 FT MOD RES 540 540 HYDROXYLATION.
 FT MOD RES 546 546 HYDROXYLATION.
 FT MOD RES 551 551 HYDROXYLATION (PARTIAL).
 FT MOD RES 552 552 HYDROXYLATION.
 FT MOD RES 561 561 HYDROXYLATION.
 FT MOD RES 603 603 HYDROXYLATION.
 FT MOD RES 610 610 IMPERFECTION IN THE GAA REPEAT.
 FT MOD RES 621 621 HYDROXYLATION (PARTIAL).
 FT MOD RES 627 627 HYDROXYLATION (PARTIAL).
 FT MOD RES 645 645 HYDROXYLATION (PARTIAL).
 FT MOD RES 647 647 HYDROXYLATION (PARTIAL).
 FT MOD RES 648 648 HYDROXYLATION.
 FT MOD RES 663 663 HYDROXYLATION.
 FT MOD RES 708 708 HYDROXYLATION.
 FT MOD RES 711 711 HYDROXYLATION.
 FT MOD RES 714 714 HYDROXYLATION.
 FT MOD RES 717 717 HYDROXYLATION.
 FT MOD RES 723 723 HYDROXYLATION.
 FT MOD RES 744 744 HYDROXYLATION.
 FT MOD RES 759 759 HYDROXYLATION.
 FT MOD RES 773 773 HYDROXYLATION.
 FT MOD RES 774 774 HYDROXYLATION (PARTIAL).
 FT MOD RES 783 783 HYDROXYLATION.
 FT MOD RES 792 792 HYDROXYLATION.
 FT MOD RES 815 815 HYDROXYLATION (PARTIAL).
 FT MOD RES 816 816 HYDROXYLATION.
 FT MOD RES 843 843 HYDROXYLATION.
 FT MOD RES 849 849 HYDROXYLATION.
 FT MOD RES 855 855 HYDROXYLATION.
 FT MOD RES 861 861 HYDROXYLATION.
 FT MOD RES 867 867 HYDROXYLATION.
 FT MOD RES 888 888 HYDROXYLATION.
 FT MOD RES 894 894 HYDROXYLATION.
 FT MOD RES 903 903 HYDROXYLATION.
 FT MOD RES 915 915 HYDROXYLATION.
 FT MOD RES 933 933 HYDROXYLATION (PARTIAL).
 FT MOD RES 939 939 HYDROXYLATION.
 FT MOD RES 945 945 HYDROXYLATION (PARTIAL).
 FT MOD RES 954 954 HYDROXYLATION.
 FT MOD RES 963 963 HYDROXYLATION.
 FT MOD RES 966 966 HYDROXYLATION.
 FT MOD RES 984 984 HYDROXYLATION.
 FT MOD RES 990 990 HYDROXYLATION (PARTIAL).
 FT MOD RES 1010 1010 HYDROXYLATION.
 FT MOD RES 1011 1011 HYDROXYLATION (PARTIAL).
 FT MOD RES 1013 1013 HYDROXYLATION.
 FT MOD RES 1014 1014 HYDROXYLATION (PARTIAL).
 FT MOD RES 1016 1016 HYDROXYLATION.
 FT MOD RES 1017 1017 HYDROXYLATION.
 FT MOD RES 1019 1019 HYDROXYLATION (PARTIAL).

CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (See http://www.isib.ch/announce/)
CC		or send an email to license@isib.ch .
CC		-----
DR	EMBL; X53556; CAA37624.1; -	
DR	PIR; S13301; S13301.	
DR	InterPro; IPR001073; Clq.	
DR	InterPro; IPR000087; Collagen.	
DR	pIam; PF00386; Clq; 1.	
DR	pIam; PF01391; Collagen; 9.	
DR	PRINTS; PR00007; COMPLEMENTC1Q.	
DR	ProDom; PD000007; Collagen; 1.	
DR	SMART; SM00110; Clq; 1.	
DR	PROSITE; PS01113; Clq; 1.	
KW	Extracellular matrix; Connective tissue; Repeat; Hydroxylation;	
KW	Cartilage; Collagen; signal; Glycoprotein.	
FT	SIGNAL	1
FT	CHAIN	18
FT	DOMAIN	19 674
FT	DOMAIN	19 56
FT	DOMAIN	57 519
FT	DOMAIN	520 674
FT	DOMAIN	539 674
FT	DISULFID	194 197
FT	MOD_RSS	460 460
FT	MOD_RSS	463 463
FT	CARBOHYD	611 611
SEQ	SEQUENCE	674 AA; 65546 MW; CD4CA73A03E004CA CRC64;
Query Match		7.8%; Score 145; DB 1; Length 674;
Best Local Similarity		24.2%; Pred. No. 0.016; Matches 100; Conservative 30; Identical 125; Indels 158; Gaps 24;
OY	15 AALLGLVGLTSLSELYVMNSANKPFAAGGYLYLPOLGDRLLDCPPRA-----65	
Dd	6 ALLILMSLNLVHG-----VFY---TERQTPTG-IKGPSTTKQFFIPAIKKGVSLR 56	
OY	66 -----RPPGP-----HSPNYEFKYLVG-----GAQR-RCEAPPAENLLLT 103	
Dd	57 GEOGIPGGPGACGPGHGHGS-----GRPKPGTGSDPGQPQLPGRPDS--A 104	
OY	104 CDRPLDLRFITKPEYSYNLMGHEFRSHNDYYLIATSDGTREGESLQG-----153	
Dd	105 TGRPL-----PLRGKO-----GER-GLNGKKGDIIGRAPLRGP 137	
OY	154 -----GCVLTRGNKYLRLVG-QSPRCGAVR-----KPYSEMP---MERDRGAHS L 196	
Dd	138 RGPFGPIPGPAGISVGRKPQCPTEBPGRGPGKEKTSYGVLGDKGEMCHCTPC 197	
OY	197 EPKKNLEPDPTSNATSGARGRLRP-----SWAYAGAAGLALLLLGVAGAGA 248	
Dd	198 RPGEGLG-----PGPTGPGEPPEVGRGENGLRGQRLKD-----QGVFERKA 245	
OY	249 MCMRRRAKPESSRHP-----GPSFGRGSGSLGGGGGMGREAEFGELIA-----296	
Dd	246 -----AGBGPQGPPEGQPEGIGKRKAGCIPOQPGIPMKKQGPAGAPLAGARGF 298	
OY	297 -----LRG-GGAADPPFCPHYENVASDDYHPITYVDGP-----POSPPRI 336	
Dd	299 GKPGLPGLKQGRGVPGLPGSPAKGEQGPAPHGEGALPGPSGNNGPGPGPKGI 351	
RESULT 34		
CAIH HUMAN		
ID	CAIH_HUMAN	STANDARD; PRT; 1516 AA.
AC	P39060; Q9Y608; Q9Y607; Q9UK38;	
DT	01-FEB-1995 (Ref. 31, Created)	
DT	16-OCT-2001 (Ref. 40, Last sequence update)	
DT	15-JUN-2002 (Ref. 41, Last annotation update)	
DE	Collagen alpha 1(XVII) chain precursor [Contains: Endostatin].	
GN	COL18A1.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	

OK NCBI_TaxID=9606;
 RN [1]
 RX SEQUENCE FROM N.A.
 RA MEDLINE=98164096; PubMed=9503365;
 RA Saarela J., Ylikarppa R., Rehn M., Purmonen S., Pihlajaniemi T.;
 RT "Complete primary structure of two variant forms of human type XVIII
 RT collagen and tissue-specific differences in the expression of the
 RT corresponding transcripts";
 RL Matrix Biol. 16:319-328 (1998).
 [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=20289799; PubMed=10830953;
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
 RA Onki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
 RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
 RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
 RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
 RA Minoshima S., Shimizu N., Kordasiek G., Hornischer K., Brandt P.,
 RA Schafte M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecher H.,
 RA Ramer J., Beck A., Klages S., Hennig S., Resselmann L., Dagand E.,
 RA Wehrhayer S., Borzym K., Gardiner K., Niketic D., Francis F.,
 RA Lehrach H., Reinhardt R., Yaspo M.-L.;
 RT "The DNA sequence of human chromosome 21.";
 RL Nature 405:311-319 (2000).
 [3]
 RN SEQUENCE OF 834-1516 FROM N.A.
 RX MEDLINE=94245237; PubMed=8188291;
 RA Oh S.-P., Warman M.L., Seidlin M.F., Cheng S., Knoll J.H., Timmons S.,
 RA Olsen B.R.;
 RT "Cloning of cDNA and genomic DNA encoding human type XVIII collagen
 RT and localization of the alpha 1(XVIII) collagen gene to mouse
 RT chromosome 19 and human chromosome 21.";
 RL Genomics 19:494-499 (1994).
 [4]
 RN SEQUENCE OF 1334-1516 FROM N.A.
 RX TISSUE=Placenta;
 RA Zhi-Yong H., Bao L., Wei-Jie Z., Xiang-Fu W.;
 RT "Cloning and expression of human endostatin gene in Escherichia
 RT coli.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 [5]
 RN INVOLVEMENT IN KNOBLOCH SYNDROME.
 RX MEDLINE=20400145; PubMed=10942434;
 RA Sertie A.L., Soresi V., Camargo A.A., Zatz M., Brähe C.,
 RA Passos-Bueno M.R.;
 RT "Collagen XVIII, containing an endogenous inhibitor of angiogenesis
 RT and tumor growth, plays a critical role in the maintenance of retinal
 RT structure and in neural tube closure.";
 RL Hum. Mol. Genet. 9:2051-2058 (2000).
 [6]
 RN VARIANT ASN-1437.
 RX MEDLINE=21518361; PubMed=11606364;
 RA Iuganelli P., Suzuki O., Godol P.H., Alves V.A., Sertie A.L.,
 RA Zorick T., Soares F., Camargo A., Moreira E.S., di Loreto C.,
 RA Moreira-Filho C.A., Simpson A., Oliva G., Passos-Bueno M.R.;
 RT "A polymorphism in endostatin, an angiogenesis inhibitor, predisposes
 RT for the development of prostatic adenocarcinoma";
 RL Cancer Res. 61:7375-7378 (2001).
 [7]
 RN FUNCTION: COL18A PROBABLY PLAYS A MAJOR ROLE IN DETERMINING THE
 CC RETINAL STRUCTURE AS WELL AS IN THE CLOSURE OF THE NEURAL TUBE.
 CC -1 FUNCTION: ENDOSTATIN POTENTIALLY INHIBITS ENDOTHELIAL CELL
 CC PROLIFERATION AND ANGIOGENESIS. MAY INHIBIT ANGIOGENESIS BY
 CC BINDING TO THE HEPARAN SULPHATE PROTEOGLYCAN INVOLVED IN GROWTH
 CC FACTOR SIGNALING.
 CC -1 ALTERNATIVE PRODUCTS. 2 ISOFORMS, A SHORT FORM/NCI-303 AND A LONG
 CC FORM/NCI-493 (SHOWN HERE), ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1 TISSUE SPECIFICITY: PRESENT IN MULTIPLE ORGANS WITH HIGHEST LEVELS
 CC IN LIVER, LUNG AND KIDNEY.
 CC -1 PTM: PROLINES ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -1 POLYMORPHISM: There is an association between a polymorphism in

CC position 1437 and prostate cancer. Heterozygous Asn-1437
 CC individuals have a 2.5 times increased chance of developing
 CC prostate cancer as compared with homozygous Asp-1437 individuals.
 CC DISEASE: Defects in COL18A1 are a cause of Knobloch syndrome (KS
 CC or KNO); an autosomal recessive disorder defined by the occurrence
 CC of high myopia, vitreoretinal degeneration with retinal
 CC detachment, macular abnormalities and occipital encephalocele.
 CC -1 SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH
 CC INTERRUPTED HELICES (FACIT) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC -----
 CC EMBL, AF018081; AAC39658.1; -;
 CC EMBL, AF018082; AAC39659.1; -;
 CC EMBL, AL163302; CAB90482.1; -;
 CC EMBL, L22548; AA51864.1; -;
 CC EMBL, AF184060; AF01310.1; ALT_INIT.
 CC HSSP; P39061; 1KOE.
 CC GlycoStatedb; B39060; -;
 CC Genew; HGNC:2195; COL18A1.
 CC MIM; 120328; -;
 CC MIM; 267750; -;
 CC InterPro; IPR000087; Collagen.
 CC InterPro; IPR001791; Laminin_G.
 CC InterPro; IPR003129; TSPN.
 CC Pfam; PF01391; Collagen; 7.
 CC Pfam; PF02210; TSPN; 1.
 CC SMART; SM00282; LamG; 1.
 CC SMART; SM00210; TSPN; 1.
 CC Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 CC Cell adhesion; Collagen; Glycoprotein; Signal; Alternative splicing;
 CC Polymorphism.
 CC SIGNAL 1 23
 CC CHAIN 24 1516
 CC CHAIN 1334 1516
 CC DOMAIN 24 516
 CC DOMAIN 517 550
 CC DOMAIN 551 560
 CC DOMAIN 561 640
 CC DOMAIN 641 664
 CC DOMAIN 665 786
 CC DOMAIN 787 809
 CC DOMAIN 810 892
 CC DOMAIN 893 906
 CC DOMAIN 907 948
 CC DOMAIN 949 961
 CC DOMAIN 962 1034
 CC DOMAIN 1035 1044
 CC DOMAIN 1045 1077
 CC DOMAIN 1078 1089
 CC DOMAIN 1090 1111
 CC DOMAIN 1112 1118
 CC DOMAIN 1119 1173
 CC DOMAIN 1174 1186
 CC DOMAIN 1187 1204
 CC DOMAIN 1205 1516
 CC CARBOHYD 68
 CC CARBOHYD 129
 CC CARBOHYD 164
 CC CARBOHYD 691
 CC CARBOHYD 1329
 CC DISULFID 1366
 CC DISULFID 1468
 CC SITE 1095
 CC VARSPIC 1097
 CC VARSPIC 180
 CC VARSPIC 215
 CC POTENTIAL.
 CC COLLAGEN ALPHA 1 (XVIII) CHAIN.
 CC ENDOSTATIN.
 CC NONHELIICAL REGION 1 (NC1).
 CC TRIPLE-HELIICAL REGION 1 (COL1).
 CC NONHELIICAL REGION 2 (NC2).
 CC TRIPLE-HELIICAL REGION 2 (COL2).
 CC NONHELIICAL REGION 3 (NC3).
 CC TRIPLE-HELIICAL REGION 3 (COL3).
 CC NONHELIICAL REGION 4 (NC4).
 CC TRIPLE-HELIICAL REGION 4 (COL4).
 CC NONHELIICAL REGION 5 (NC5).
 CC TRIPLE-HELIICAL REGION 5 (COL5).
 CC NONHELIICAL REGION 6 (NC6).
 CC TRIPLE-HELIICAL REGION 6 (COL6).
 CC NONHELIICAL REGION 7 (NC7).
 CC TRIPLE-HELIICAL REGION 7 (COL7).
 CC NONHELIICAL REGION 8 (NC8).
 CC TRIPLE-HELIICAL REGION 8 (COL8).
 CC NONHELIICAL REGION 9 (NC9).
 CC TRIPLE-HELIICAL REGION 9 (COL9).
 CC NONHELIICAL REGION 10 (NC10).
 CC TRIPLE-HELIICAL REGION 10 (COL10).
 CC NONHELIICAL REGION 11 (NC11).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC O-LINKED (GALNAC. . .) (POTENTIAL).
 CC O-LINKED (GALNAC. . .) (POTENTIAL).
 CC /FTID-CAR 000150.
 CC BY SIMILARITY.
 CC CELL ATTACHMENT SITE (POTENTIAL).
 CC MISSING (IN SHORT ISOREM).
 CC HTTBACTLPAPTPSPSLGRPAAPALPGSPVPPSS -> MA

[illegible]

```

RT      "Cloning and developmental expression of the alpha 3 chain of chicken
RT      type IX collagen."
RL      J. Biol. Chem. 267:10070-10076 (1992).
CC      -1- FUNCTION: COLLAGEN TYPE IX IS A MINOR CARTILAGE NONFIBRILLAR
CC      COLLAGEN. IT IS ASSOCIATED WITH TYPE II COLLAGEN FIBRILS.
CC      -1- SUBUNIT: TRIMERS COMPOSED OF THREE DIFFERENT CHAINS: ALPHA 1 (IX) ,
CC      ALPHA 2 (IX), AND ALPHA 3 (IX).
CC      -1- PMM: PROLINS ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC      UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi.ac.uk/announcement/
CC      or send an email to license@ebi.ac.uk).
CC      -----
DR      EMBL; M83179; AAB59960.1; -.
DR      PIR; S22918; S22918.
DR      InterPro; IPR000087; Collagen.
DR      Pfam; PF01391; Collagen; 11.
DR      ProDom; PD000007; Collagen; 1.
KW      Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW      Glycoprotein; Cartilage; Collagen; Signal.
FT      SIGNAL          1
FT      CHAIN          21
FT      DOMAIN         22      675
FT      DOMAIN         25      515
FT      DOMAIN         516      546
FT      DOMAIN         547      626
FT      DOMAIN         627      631
FT      DOMAIN         632      658
FT      DOMAIN         659      675
FT      SITE           242      244
FT      SITE           591      593
FT      CARBOHYD       479      479
FT      SEQUENCE       675 AA; 63013 MW; C9831BC924A10098 CRC64;
SQ
Query Match      7.8%; Score 144; DB 1; Length 675;
Best Local Similarity 28.8%; Pred. No. 0.018;
Matches 64; Conservative 9; Mismatches 71; Indels 78; Gaps 12;
DB 166 RVGGSPRGCAVPRKRVSMERBDGAHSLPEPKENLPDPTSNATSRGAGPLPPPSM 225
DB 23 RVG-PPGPRGPRGSGKDIDG-----EPGSGLPGRP-----GPKGAGX 65
DB 226 PAVAGAG-GLATLLGVAGAGAMCWRRRRANKPSSRHG-----PGFGRG- 272
DB 66 PQAAGEAGLPGLP-----GVDPGLTG-----DPPGPRGPPGDRGALGPAGPPGAGKL 115
DB 273 -----GSLGGGGGG-----GPR-----EAEPGELGALRG----- 300
DB 116 PGPSPPPSPGLPGANGRGPRPGSLGFPGRPPPPPLGALGILPPGGGLQCPALCP 175
DB 301 -GAADPFCPHYEKVSGDYGHPIVIVQDP-----PQSPNNI 336
DB 176 PGPSPGMPGPKGHTGKGGGGRIGKEGKSPGPPPPGI 217
RESULT 36
CAL9_HUMAN          STANDARD; PRT; 921 AA.
AC      P20849; Q9Y6P2; Q9Y6P3; Q9H151; Q9H152; Q99225; Q13699; Q13700;
DT      01-FEB-1991 (Rel. 17, Created)
DT      15-JUN-2002 (Rel. 41, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Collagen alpha 1 (IX) chain precursor.
GN      COL9A1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_Taxid=9606;
RN      [1]

```


Query Match 7.8%; Score 144; DB 1; Length 921;
 Best Local Similarity 22.3%; Pred. No. 0.025;
 Matches 84; Conservative 21; Mismatches 120; Indels 152; Gaps 13;

CC 2 GPHSGPGVRRGALLLGLVGLSLPEPYMSANKRQAEQGVYLPQIGRDL 61
 DB 391 GPP--GPPGPR-----GTIGFHDG-----DPL 410

CC 62 CRRAPPPGPHSSPNVEFYKLYIVG-----GAQRRRCAPAPNLLITCDRDLRFTIK 116
 DB 411 CRRAPPPGPHSSPNVEFYKLYIVG-----GAQRRRCAPAPNLLITCDRDLRFTIK 116

CC 117 PCEYSPNLMGHEFRSHHDYIATSDGTREGESLQGVCLTRGKVLRYGQSPRGAV 176
 DB 441 -----PGRGR-----KGEEDQG-----ELGEGAGQPPGQ 468

CC 177 PRKPYSEMPMERDRGAHSL--EPGKENVPGDPTSNAT-----SRGAGPLEPP 223
 DB 469 GURGITGVGDKGKAGLDGEPQGLPGAPDQGGORPGEPGAGPGRGAGAGKIF 528

CC 224 SMPAVAGAGGALLLGVAGAGAMCRRRRAKSESRRHG-----PSFRGSLG 276
 DB 529 GLPGRGDTG-----LPVDGRDGI.PGMPGTGEPKPGPPEDAGLQGLPEVPGIPGAKG 583

CC 277 LGGGGGMPREAPGELGIALRG-----GAADPFCPHYEKVSG 316
 DB 584 VAGEKGSITAPKPGQMGNSGKPGQGGPPGEGVPRPGQLPQSRRELTPVSGPLPKLG 643

CC 317 DYGHPIYIVQDGPPOSP 333
 DB 644 SLGSPGLPGLPGPPLP 660

RESULT 37
 CHICK CHICK STANDARD; PRT; 674 AA.

AC P08125;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Collagen.alpha.1(X) chain precursor.
 GN COL10A1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN 111
 RP SEQUENCE OF 46-674 FROM N.A., AND SEQUENCE OF 103-117 AND 453-466.
 RX MEDLINE=86168227; PubMed=3082876;
 RA Nimmiya Y., Gordon M., van der Rest M., Schmid T., Linsemayer T.,
 Olsen B.R.;
 RT "The developmentally regulated type X collagen gene contains a long
 RT open reading frame without introns."
 RL J. Biol. Chem. 261:5041-5050(1986).
 RN 112
 RP SEQUENCE OF 1-75 FROM N.A.
 RX MEDLINE=89054019; PubMed=2461368;
 RA Luvall P., Nimmiya Y., Rosenblum N.D., Olsen B.R.;
 RT "The type X collagen gene. Intron sequences split the 5'-untranslated
 RT region and separate the coding regions for the non-collagenous amino-
 RT terminal and triple-helical domains."
 RL J. Biol. Chem. 263:18378-18385(1988).
 RN 113
 RP REVISIONS TO C-TERMINUS.
 RX MEDLINE=89380199; PubMed=2476437;
 RA Yamasuchi N., Benya P.D., van der Rest M., Nimmiya Y.;
 RT "The cloning and sequencing of alpha 1(VIII) collagen cDNAs
 RT demonstrate that type VIII collagen is a short chain collagen and
 RT contains triple-helical and carboxyl-terminal non-triple-helical
 RT domains similar to those of type X collagen."
 RL J. Biol. Chem. 264:16022-16029(1989).
 CC -!- FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTROPHIC

CC CHONDROCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE
 CC MINERALIZATION ZONES OF HYALINE CARTILAGE.
 CC -!- SUBUNIT: HOMOTRIMER.
 CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -!- SIMILARITY: STRONG, TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.
 CC -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC
 CC EMBL; M13496; AAA48736.1; ALT_SEQ.
 CC EMBL; J04194; AAA48634.1; -.
 CC PIR; A31896; A31896.
 CC InterPro; IPR001073; C1Q.
 CC InterPro; IPR000087; Collagen.
 CC Pfam; PF01391; C1Q; 1.
 CC Pfam; PF01391; Collagen; 8.
 CC PRINTS; PR00007; COMPLEMENTC1Q.
 CC SMART; SM00110; C1Q; 1.
 CC PROSITE; PS01113; C1Q; 1.
 CC DR Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 CC KW Cartilage; Collagen; Signal.
 CC FT SIGNAL 1 18
 CC FT CHAIN 19 674 COLLAGEN ALPHA 1(X) CHAIN.
 CC FT DOMAIN 19 52 NONHELICAL REGION (NC2).
 CC FT DOMAIN 53 512 TRIPLE-HELICAL REGION.
 CC FT DOMAIN 513 674 NONHELICAL REGION (NC1).
 CC FT DOMAIN 539 674 C1Q.
 CC FT MOD_RES 453 453 HYDROXYLATION.
 CC FT MOD_RES 456 456 HYDROXYLATION.
 CC FT MOD_RES 456 456
 CC SQ SEQUENCE 674 AA; 66434 MW; EAB48B1EF74B145 CRC64;

Query Match 7.7%; Score 143; DB 1; Length 674;
 Best Local Similarity 25.0%; Pred. No. 0.021;
 Matches 77; Conservative 18; Mismatches 91; Indels 122; Gaps 16;

CC 67 PRGPHSSPNVEFY-KLYIVG-----GAQRRRCAPAPNLLITCDRDLRFTIK 117
 DB 86 PRGPHSSPNVEFY-KLYIVG-----GAQRRRCAPAPNLLITCDRDLRFTIK 117

CC 118 QEYSPNLMGHEFRSHHDYIATSDGTREGESLQGVCLTRGKVLRYGQSPRGAV 171
 DB 126 -----PGRGR-----KGEEDQG-----ELGEGAGQPPGQ 468

CC 172 RGAAPPRKPYSEMPMERDRGAHSLPEPKENVPGDPTSNAT-----SRGAG 218
 DB 157 RGAAPPRKPYSEMPMERDRGAHSLPEPKENVPGDPTSNAT-----SRGAG 218

CC 219 PLRPPSPAVAGAGGALLLGVAGAGAMCRRRRAKSESRRHG-----PSFRGSLG 276
 DB 200 PLRPPSPAVAGAGGALLLGVAGAGAMCRRRRAKSESRRHG-----PSFRGSLG 276

CC 275 LGGGGGMPREAPGELGIALRG-----GAADPFCPHYEKVSG 316
 DB 254 LGGGGGMPREAPGELGIALRG-----GAADPFCPHYEKVSG 316

CC 317 DYGHPIYIVQDGPPOSP 333
 DB 306 DYGHPIYIVQDGPPOSP 333

RESULT 38
 CA25 HUMAN STANDARD; PRT; 1496 AA.

AC P05957;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 15-JUN-2002 (Ref. 41, last annotation update)
DE Collagen alpha 2(V) chain precursor.
CN COL5A2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
R1 SEQUENCE OF 1-463 FROM N.A.
RX MEDLINE=69123368; PubMed=2914927;
RA Woodbury D., Benson-Chanda V., Ramirez F.,
RT "Amino-terminal propeptide of human pro-alpha 2(V) collagen conforms
RL to the structural criteria of a fibrillar procollagen molecule.";
RL J. Biol. Chem. 264:2735-2736(1989).
RN [2]
R1 SEQUENCE OF 398-1496 FROM N.A.
RX MEDLINE=6714633; PubMed=3029669;
RA Well D., Bernard M.P., Gargano S., Ramirez F.,
RT "The pro alpha 2(V) collagen gene is evolutionarily related to the
RL major fibrillar-forming collagens";
RL Nucleic Acids Res. 15:181-198(1987).
RN [3]
R1 SEQUENCE OF 1227-1496 FROM N.A.
RX MEDLINE=8528933; PubMed=2411731;
RA Myers J.C., Loidl H.R., Seyer J.M., Dion A.S.,
RT "Complete primary structure of the human alpha 2 type V procollagen
RL COOH-terminal propeptide.";
RL J. Biol. Chem. 260:11216-11222(1985).
RN [4]
R1 SEQUENCE OF 1448-1496 FROM N.A.
RX MEDLINE=89138450; PubMed=3224983;
RA Tsipouras P., Schwartz R.C., Liddell A.C., Salveid C.S., Well D.,
RL Ramirez F.,
RT "Genetic distance of two fibrillar collagen loci, COL3A1 and COL5A2,
RL located on the long arm of human chromosome 2.";
RL Genomics 3:275-277(1988).
RN [5]
R1 SEQUENCE OF 208-227.
RX TISSUE=Placenta;
RL MEDLINE=92239022; PubMed=1571108;
RA Mann K.,
RT "Isolation of the alpha 3-chain of human type V collagen and
RL characterization by partial sequencing.";
RL Biol. Chem. Hoppe-Seyler 373:69-75(1992).
RN [6]
R1 SEQUENCE OF 288-297 AND 606-617.
RX TISSUE=Bone;
RL MEDLINE=9423164; PubMed=8181482;
RA Moradl-Anelli M., Rousseau J.C., Klamann J.P., Champilaud M.F.,
RL Bouillon M.M., Bernillon J., Wallach J.M., van der Rest M.,
RT "Diversity in the processing events at the N-terminus of type-V
RL collagen.";
RL Eur. J. Biochem. 221:987-995(1994).
RN [7]
R1 DISEASE.
RX PubMed=9425231;
RA Michalichova K., Susic M., Willing M.C., Wenstrup R.J., Cole W.G.,
RT "Mutations of the alpha2(V) chain of type V collagen impair matrix
RL assembly and produce Ehlers-Danlos syndrome type I.";
RL Hum. Mol. Genet. 7:249-255(1998).
RN [8]
R1 VARIANT EDS-II ARG-960.
RX PubMed=9783710;
RA Richards A.J., Martin S., Nicholas A.C., Harrison J.B., Pope F.M.,
RL Burrows N.P.,
RT "A single base mutation in COL5A2 causes Ehlers-Danlos syndrome type
RL II.";
RL J. Med. Genet. 35:846-848(1998).
RN [9]
R1 FUNCTION: TYPE V COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
RX (FIBRILLAR FORMING COLLAGEN). IT IS A MINOR CONNECTIVE TISSUE
CC COMPONENT OF NEARLY UBQUITOUS DISTRIBUTION. TYPE V COLLAGEN BINDS
CC TO DNA, HEPARAN SULFATE, THROMBOSPONDIN, HEPARIN, AND INSULIN.
CC -1 SUBUNIT: TRIMERS OF TWO ALPHA 1(V) AND ONE ALPHA 2(V) CHAINS IN

MOST TISSUES AND TRIMERS OF ONE ALPHA 1(V), ONE ALPHA 2(V), AND
ONE ALPHA 3(V) CHAINS IN PLACENTA.

-I- PFM; PROLINS AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.

-I- DISEASE: Defects in COL5A2 are a cause of Ehlers-Danlos syndrome gravis or type I (EDS-I), and Ehlers-Danlos syndrome mitis or type II (EDS-II). Ehlers-Danlos syndrome is a genetically and phenotypically heterogeneous connective-tissue disorder characterized by loose-jointedness and fragile, velvety, stretchable, bruisable skin that heals with peculiar "cigarette-paper" scars. EDS-I and EDS-II are autosomal dominant traits.

-I- SIMILARITY: CONTAINS 1 WVF C DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@lsb.sib.ch). -----

EMBL; J04478; AA51859.1; -.
DR EMBL; X04758; CA28454.1; -.
DR EMBL; M11718; AA52058.1; -.
DR PIR; A25374; A25374.
DR PIR; A25874; A25874.
DR PIR; A30017; A30017.
DR PIR; A31427; A31427.
GeneW; HGNC:2210; COL5A2.
MIM; 120190; -.
MIM; 130000; -.
DR MIM; 130010; -.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000885; Fib collagen_C.
DR InterPro; IPR001007; WVF_C.
Pfam; PF000093; WVC; 1.
PFam; PF01391; Collagen; 18.
PFam; PF01410; COLF1; 1.
ProdDom; PD000007; Collagen; 5.
Dr Prodom; PD002078; Fib collagen_C; 1.
SMART; SMO0038; COLFI; 1.
SMART; SMO0214; WMC; 1.
PROSITE; PS01208; WMFC; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Collagen; Signal; Ehlers-Danos syndrome;
KW Disease mutation.
FT SIGNAL 1 26
FT CHAIN 27 1226 COLLAGEN ALPHA 2(V) CHAIN.
FT PROEP 1227 1496 CARBOXYL-TERMINAL PEPTIDE.
FT DOMAIN 39 97 WVCF.
FT MOD_RES 290 290 HYDROXYLATION.
FT MOD_RES 293 293 HYDROXYLATION.
FT MOD_RES 296 296 HYDROXYLATION.
FT MOD_RES 608 608 HYDROXYLATION.
FT MOD_RES 614 614 G-> R (IN EDS-II).
FT VARIANT 960 960 /FITD-VAR_013588.
 A -> P (IN REF. 6).
FT CONFLICT 292 292 K -> T (IN REF. 3).
FT CONFLICT 1418 1418 F -> Q (IN REF. 3).
FT CONFLICT 1438 1438 E -> S (IN REF. 4).
FT CONFLICT 1460 1460 V -> A (IN REF. 4).
FT CONFLICT 1496 1496 V -> A (IN REF. 4).
SQ SEQUENCE 1496 AA; 144720 MW; 82827C17A8644F5A CRC64;

Query Match 7 7%; Score 142.5; DB 1; Length 1496;
Beet Local Similarity 29.1%; Pred. No. 0.051; Indels 43; Gaps 9;
Matches 55; Conservative 15; Mismatches 76;

Gy 169 GSFRG--GA-V-FRRYVSMPWRDRGAHSHLEPKENVPG-----DPTSNTSRGAE 217
Db 500 RGPRGDDETLGRPPGVVRGARFGNRF-----PSDDGPGRKAGQRGVPVGSGGGPSQS 554
Gy 218 GPUPPPSPNVAVGAAGLLALLLVAGAAGAMCMRRRAKSESERHGCP----GSFGRG 273

```

Db 555 GDBRGPRGPRGLRGARG-----LTGNPVGQSPBGKGLGALGAGDEDRPFPSPSITKQPG 609
QY 274 SLGLGGGGGCGRRAPGELGLALRGGAADPPFCPHYKVSQDYGHPVTVQGGP--- 330
Db 610 TWGLPGPGKSNQDPKPGPAG---NPGVPGQGAFAFGKHGVPYGPFP-----GPGQLR 659
QY 331 -----QSPG 334
Db 660 GRRGQGP 668

RESULT 39
CA18 HUMAN STANDARD; PRT; 744 AA.
AC P2758; 096D07;
DT 01-AUG-1992 (Rel. 23, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Collagen alpha 1(VIII) chain precursor (Endothelial collagen).
GN COL8A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=91231001; PubMed=2029894;
RA Muragaki Y., Mattei M.-G., Yamaguchi N., Olsen B.R., Nicomly Y.;
RT "The complete primary structure of the human alpha 1 (VIII) chain and
RT assignment of its gene (COL8A1) to chromosome 3";
RL Eur. J. Biochem. 197; 615-622(1991).
RN 12
RP SEQUENCE FROM N.A.
RC Tissue=Lung;
RA Straube R.;
RL Submitted (Sep-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAJOR COMPONENT OF THE DESCENDING MEMBRANE (BASEMENT
CC MEMBRANE) OF CORNEAL ENDOTHELIAL CELLS.
CC -1- SUBUNIT: MAY FORM HOMOTRIMERS, OR HETEROTRIMERS IN ASSOCIATION
CC WITH ALPHA 2(VIII) TYPE COLLAGENS.
CC -1- PM: PROLINES AT THE THIRD POSITION OF THE TRIPLET REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- MISCELLANEOUS: FOUR CONSECUTIVE GLY-PRO-PRO TRIPLETS ARE PRESENT
CC AT THE C-TERMINUS OF THE TRIPLE-HELICAL REGION. THESE MAY PROVIDE
CC THE HIGH THERMAL STABILITY OF THIS REGION.
CC -1- SIMILARITY: STRONG, TO ALPHA 2 TYPES VIII AND X COLLAGENS.
CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/submit/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; X57527; CAA40748.1; -
CC EMBL; BC013581; AAH13581.1; -
CC PIR; S15435; S15435.
CC Genew; HGNC:2215; COL8A1.
CC MIM; 120251; -
CC InterPro; IPR001073; C1Q.
CC InterPro; IPR000877; Collagen.
CC Pfam; PF00386; C1Q; 1.
CC Pfam; PF00391; Collagen_8.
CC PRINTS; PR00007; COMPLEMENTC1Q.
CC ProDom; PD00007; Collagen; 1.
CC SMART; SM00110; C1Q; 1.
CC PROSITE; PS01113; C1Q; 1.
DR Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Cell adhesion; Collagen; Signal.
FT SIGNAL 1 28

```

```

FT CHAIN 29 744 COLLAGEN ALPHA 1(VIII) CHAIN.
FT DOMAIN 29 117 NONHELICAL REGION (NC2).
FT DOMAIN 118 571 TRIPLE-HELICAL REGION (COL1).
FT DOMAIN 572 744 NONHELICAL REGION (NC1).
FT C1Q.
FT DOMAIN 609 744
FT CONFLICT 262 262 P -> L (IN REF. 1).
FT CONFLICT 297 297 P -> R (IN REF. 1).
FT CONFLICT 344 344 P -> A (IN REF. 1).
FT CONFLICT 382 382 A -> S (IN REF. 1).
FT CONFLICT 388 388 P -> S (IN REF. 1).
FT CONFLICT 454 454 L -> F (IN REF. 1).
FT CONFLICT 464 464 A -> H (IN REF. 1).
FT CONFLICT 601 601 Y -> T (IN REF. 1).
FT CONFLICT 631 631 A -> G (IN REF. 1).
SQ SEQUENCE 744 AA; 73364 MW; 28C1B0955DE2C9A3 CRC64;

Query Match 7.7%; Score 142; DB 1; Length 744;
Best Local Similarity 23.3%; Pred. No. 0.027;
Matches 72; Conservative 16; Mismatches 99; Indels 122; Gaps 13;

QY 67 PPGHSSPNYEFYKLYLVGAQGRCEAPAPNLLLTCDRPLDLRFITKFGESPNLWG 126
Db 198 PPGPHSLPG-----TKPGGPGGLPGQPGK----- 222
QY 127 HEFRSHHDYIATSDGTREGLESQGVCLTRGKVKLLRVQSPRGAVPRKVSMPM 186
Db 223 -----GDRGPKGLPGPGQ-----LRGPKDKGFGMPGAPGVKGP- 256
QY 187 ERDGAASLEPKENTLPDPTSNMATS-RGAGPLPSPMPAVAAAGLALLL-----G 241
Db 257 ---PQNHG-PGPGVGLPEVGVKPGVTPGPGQPLGKAPGEPGPGTIGVPGVQPGP 311
QY 242 VAGAGGAMCMRRRRAPKPSERHPG-----PGSRGGS 274
Db 312 IPGIG-----KPGDGI PGQPGPGGKGEGGLPGLPGLPGIGKPGPGPKD 361
QY 275 LGIGG-GGGMGRRAPGELGLALRGGAADPPFCPHYKVSQDYGHPVTVQDGP----- 329
Db 362 RKGGVPGALGPR-GEKPIDGPIGIG-----PPGPGPLGPIGPGFPAIGFPGKGG 416
QY 330 -----QSPG 334
Db 417 GIVPGGP 425

RESULT 40
CA26 MOUSE STANDARD; PRT; 1029 AA.
AC 002788; 005505;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Collagen alpha 2(VI) chain precursor.
GN COL6A2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RA Ibrahim A., Bardon S., Dani C.;
RL Submitted (May-1992) to the EMBL/GenBank/DBJ databases.
CC [2]
CC SEQUENCE OF 266-1029 FROM N.A.
CC MEDLINE=93143659; PubMed=8380980;
CC Ibrahim A., Bertrand B., Bardon S., Amri E.Z., Grimaldi P.,
CC Ailhaud G., Dani C.;
CC "Cloning of alpha 2 chain of type VI collagen and expression during
CC mouse development.";
CC Biochem. J. 289:141-147(1993).
RN 13
RN SEQUENCE OF 266-600 FROM N.A.
RN STRAIN=C57BL/6; TISSUE=Fibroblast;
RC

```

RX MEDLINE=91226374; PubMed=1709252;
 RA Constantino C.D., Jimenez S.A.;
 RT "Structure of CDNA's encoding the triple-helical domain of murine
 RT alpha 2 (VI) collagen chain and comparison to human and chick
 RT homologues. Use of polymerase chain reaction and partially degenerate
 RT oligonucleotide for generation of novel cDNA clones.";
 RL Matrix 11:1-9(1991).
 [4]
 RP SEQUENCE OF 659-1029 FROM N.A.
 RX MEDLINE=93256888; PubMed=8469506;
 RA Zhang R.Z., Pan T.C., Timpl R., Chu M.-L.;
 RT "Cloning and sequence analysis of cDNAs encoding the alpha 1, alpha 2
 RT and alpha 3 chains of mouse collagen VI.";
 RL Biochem. J. 291:787-792(1993).
 CC -1- FUNCTION: COLLAGEN VI ACTS AS A CELL-BINDING PROTEIN.
 CC -1- SUBUNIT: TRIMERS COMPOSED OF THREE DIFFERENT CHAINS: ALPHA 1 (VI),
 CC ALPHA 2 (VI), AND ALPHA 3 (VI).
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ADIPOSE TISSUE, LUNG,
 CC ADRENAL GLANDS AND OVARY. LOWER LEVELS IN TESTIS, TONGUE, SKIN,
 CC KIDNEY, HEART, INTESTINE AND SPLEEN. NO EXPRESSION IN SKELETAL
 CC MUSCLE OR LIVER.
 CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 CC -1- UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -1- SIMILARITY: CONTAINS 3 VMPA DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X65582; CAA46541.1; -
 DR EMBL; X62332; CAA44206.1; -
 DR EMBL; Z18272; CAA79153.1; -
 DR EMBL; L06343; AAA7441.1; -
 DR PIR; S13745; S13745.
 DR MGD; MG1:88460; Col6a2.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR002035; VMP_A.
 DR Pfam; PF01381; Collagen; 4.
 DR Pfam; PF00453; VMPADOMAIN.
 DR SMART; SM00327; VMA; 3.
 DR PROSITE; PS50234; VMPA; 3.
 KM Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KM Glycoprotein; Cell adhesion; Collagen; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 1029
 FT DOMAIN 21 265
 FT DOMAIN 266 600
 FT DOMAIN 601 1029
 FT DOMAIN 54 178
 FT DOMAIN 623 748
 FT DOMAIN 843 965
 FT SITE 376 378
 FT SITE 436 438
 FT SITE 499 501
 FT SITE 508 510
 FT SITE 549 551
 FT CARBOHYD 150 150
 FT CARBOHYD 337 337
 FT CARBOHYD 640 640
 FT CARBOHYD 795 795
 FT CARBOHYD 907 907
 FT CARBOHYD 963 963
 FT CONFLICT 268 268
 FT CONFLICT 804 804
 FT CONFLICT 848 848
 FT CONFLICT 962 966
 FT CONFLICT 976 977
 SQ SEQUENCE 1029 AA; 109811 MW; EDALBA0AF266FSD CRC64;

Query Match 7.78; Score 142; DB 1; Length 1029;
 Best Local Similarity 29.78; Pred. No. 0.038;
 Matches 90; Conservative 16; Mismatches 125; Indels 72; Gaps 20;
 QY 68 PGPSPVYEEYKLYLVGAGRCCEAPAPNLLTCDPRLDRTFTKQVEYFNLMGH 127
 DB 265 PGPSPVYEEYKLYLVGAGRCCEAPAPNLLTCDPRLDRTFTKQVEYFNLMGH 318
 QY 128 --EPSSHHDYITITSQGTREGLSLOG--GVCITRKQVLL-RVG-QSPRGAVPRKP- 180
 DB 319 KGEF-----GSDG-RKAGGLGAKNG--TDCQKGLGIGPGGCKGDPGRSPD 364
 QY 181 ----VSEMPWER-DRGA-ASHLEPGKENVLPDDPTSNATSGAEGFLPPSPAPVAGAAG 234
 DB 365 GYPEAGPGRGQGAKGDSGRGRGRGPPDP--GDGSKGYQANNAPGSPYKGGKGG 423
 QY 235 LALLLVGAGGAMCWRRRRAKP-----SESHHPG---GSFGRGSLGLG 278
 DB 424 -----PGRGPKEPGRRGDPGTGKGPSPDGPKEKDPGEGPGLAGEVSGKAK 475
 QY 279 GGGGW-GPRELPEGLIALRGGAADPPCPHYEKVSGYGHVYIVODGPPSP--PN 335
 DB 476 GDRGPGRGPQ-GALBPGKQSGRGP-----GAPGRGDSGQGPCKGDPGRPG 524
 QY 336 IYV 338
 DB 525 FSX 527

RESULT 41
 CA24_ASCSU STANDARD; PRT; 1763 AA.

AC P27393;
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DB Collagen alpha 2 (IV) chain precursor.
 OS Ascaris suum (pig roundworm) (Ascaris lumbricoides).
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascarididae;
 CC NCB1_TaxId=6253;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=91340768; PubMed=1714907;
 RA Pettit J., Kingston I.B.;
 RT "The complete primary structure of a nematode alpha 2(IV) collagen
 RT and the partial structural organization of its gene.";
 RL J. Biol. Chem. 266:16149-16156(1991).
 CC -1- FUNCTION: COLLAGEN TYPE IV IS SPECIFIC FOR BASEMENT MEMBRANES.
 CC -1- SUBUNIT: TRIMERS OF TWO ALPHA 1(IV) AND ONE ALPHA 2(IV) CHAIN.
 CC TYPE IV COLLAGEN FORMS A MESH-LIKE NETWORK LINKED THROUGH
 CC INTERMOLECULAR INTERACTIONS BETWEEN 7S DOMAINS AND BETWEEN NCI
 CC DOMAINS.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; I (SHOWN HERE) AND II; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
 CC DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
 CC G-X-Y REPEATS IN THE LONG CENTRAL, TRIPLE-HELICAL DOMAIN (WHICH MAY
 CC CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
 CC TRIPLE-HELICAL 7S DOMAIN.
 CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC ARE INVOLVED IN INTER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF
 CC THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE
 CC IV COLLAGENS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

RT "Covalent structure of collagen: amino acid sequence of alpha
RT (I(III)-OB9 from type III collagen of human liver.";
RL Biochemistry 20:2621-2627(1981).
RN [111]
RP SEQUENCE OF 1176-1466 FROM N.A.
RX MEDLINE=85157600; PubMed=2579949;
RA Chu M.-L., Weil D., de Wet M.J., Bernard M.P., Sippola M., Ramirez F.;
RT "Isolation of cDNA and genomic clones encoding human pro-alpha 1
RT (III) collagen. Partial characterization of the 3' end region of the
RT gene.";
RL J. Biol. Chem. 260:4357-4363(1985).
RN [12]
RP SEQUENCE OF 1161-1200 FROM N.A.
RX MEDLINE=86187804; PubMed=3754452;
RA Miskulin P., Dalgleish R., Kluge-Beckerman B., Remard S.I.,
RA Tolstoshev P., Brantly M., Crystal R.G.;
RT "Human type III collagen gene expression is coordinately modulated
RT with the type I collagen genes during fibroblast growth.";
RL Biochemistry 25:1408-1413(1986).
RN [13]
RP SEQUENCE OF 1-170 FROM N.A.
RC TISSUE-Placenta;
RX MEDLINE=88303360; PubMed=3405773;
RA Toman D., Ricca G., de Crombrughe B.;
RT "Nucleotide sequence of a cDNA coding for the amino-terminal region
RT of human prepro alpha 1(III) collagen.";
RL Nucleic Acids Res. 16:7201-7201(1988).
RN [14]
RP SEQUENCE OF 1-176 FROM N.A.
RX MEDLINE=89378752; PubMed=2777083;
RA Benson-Chanda V., Su M.W., Weil D., Chu M.-L., Ramirez F.;
RT "Cloning and analysis of the 5' portion of the human type-III
RT procollagen gene (COL3A1).";
RL Gene 78:255-265(1989).
RN [15]
RP REVIEW ON VARIANTS.
RX MEDLINE=97255959; PubMed=9101290;
RA Kuivaniemi H., Tromp G., Prockop D.J.;
RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-
RT associated collagen (type IX), and network-forming collagen (type X)
RT cause a spectrum of diseases of bone, cartilage, and blood vessels.";
RL Hum. Mutat. 9:300-315(1997).
RN [16]
RP VARIANT AORTIC ANEURYSM ARG-303, AND VARIANT THR-668.
RX MEDLINE=93393988; PubMed=8514866;
RA Tromp G., Wu Y., Prockop D.J., Madhathari S.L., Kleiner C.,
RA Barley J.J., Zhuang J., Noerregaard O., Darling R.C., Abbott W.M.,
RA Cole C.W., Jaakkola P., Rytanen M., Pearce W.H., Yao J.S.T.,
RA Majamaa K., Michels S.N., Gatalica Z., Ferrelli R.B., Jimenez S.A.,
RA Jackson C.E., Smillens V.V., Kaya M., Kuivaniemi H.;
RT "Sequencing of cDNA from 50 unrelated patients reveals that mutations
RT in the triple-helical domain of type III procollagen are an
RT infrequent cause of aortic aneurysms.";
RL J. Clin. Invest. 91:2539-2545(1993).
RN [17]
RP VARIANT THR-698.
RX MEDLINE=91045136; PubMed=2235526;
RA Zafarullah K., Kleiner C., Tromp G., Kuivaniemi H., Kontusaari S.,
RA Wu Y., Ganguly A., Prockop D.J.;
RT "A mutation in the gene for type III procollagen (COL3A1) in a family
RT with aortic aneurysms.";
RL J. Clin. Invest. 86:1465-1473(1990).
RN [18]
RP VARIANT AORTIC ANEURYSM ARG-786.
RX MEDLINE=91056145; PubMed=2243125;
RA Kontusaari S., Tromp G., Kuivaniemi H., Romanic A.M., Prockop D.J.;
RT "A mutation in the gene for type III procollagen (COL3A1) in a family
RT with aortic aneurysms.";
RL J. Clin. Invest. 86:1465-1473(1990).
RN [19]
RP VARIANT EDS-IV ARG-828.
RX MEDLINE=94016385; PubMed=8411057;
RA Richards A.J., Narcisi P., Lloyd J.C., Ferguson C., Pope F.M.;
RT "The substitution of glycine 661 by arginine in type III collagen

RT produces mutant molecules with different thermal stabilities and
RT causes Ehlers-Danlos syndrome type IV.";
RL J. Med. Genet. 30:690-693(1993).
RN [20]
RP VARIANT EDS-IV SER-957.
RX MEDLINE=89109135; PubMed=2492273;
RA Tromp G., Kuivaniemi H., Shikata H., Prockop D.J.;
RT "A single base mutation that substitutes serine for glycine 790 of
RT the alpha 1 (III) chain of type III procollagen exposes an arginine
RT and causes Ehlers-Danlos syndrome IV.";
RL J. Biol. Chem. 264:1349-1352(1989).
RN [21]
RP VARIANT EDS-IV VAL-960.
RX MEDLINE=85266429; PubMed=7749417;
RA Tromp G., de Paepe A., Nuytink L., Madhathari S.L., Kuivaniemi H.;
RT "Substitution of valine for glycine 793 in type III procollagen in
RT Ehlers-Danlos syndrome type IV.";
RL Hum. Mutat. 5:179-181(1995).
RN [22]
RP VARIANT EDS-IV GLU-1014.
RX MEDLINE=92316511; PubMed=1352273;
RA Richards A.V., Ward P.N., Narcisi P., Nicholls A.C., Lloyd J.C.,
RA Pope F.M.;
RT "A single base mutation in the gene for type III collagen (COL3A1)
RT converts glycine 847 to glutamic acid in a family with Ehlers-Danlos
RT syndrome type IV. An unaffected family member is mosaic for the
RT mutation.";
RL Hum. Genet. 89:414-418(1992).
RN [23]
RP VARIANT EDS-IV ASP-1050.
RX MEDLINE=90037070; PubMed=2808425;
RA Tromp G., Kuivaniemi H., Stolle C.A., Pope F.M., Prockop D.J.;
RT "Single base mutation in the type III procollagen gene that converts
RT the codon for glycine 883 to aspartate in a mild variant of
RT Ehlers-Danlos syndrome IV.";
RL J. Biol. Chem. 264:19313-19317(1989).
RN [24]
RP VARIANT EDS-IV VAL-1077.
RX MEDLINE=91374480; PubMed=1895316;
RA Richards A.J., Lloyd J.C., Ward P.N., de Paepe A., Narcisi P.,
RA Pope F.M.;
RT "Characterization of a glycine to valine substitution at amino acid
RT position 910 of the triple helical region of type III collagen in a
RT patient with Ehlers-Danlos syndrome type IV.";
RL J. Med. Genet. 28:458-463(1991).
RN [25]
RP VARIANT EDS-IV GLU-1173.
RX MEDLINE=93022543; PubMed=1357232;
RA Johnson P.H., Richards A.J., Pope F.M., Hopkinson D.A.;
RT Query Match 7.6%; Score 141.5; DB 1; Length 1466;
RT Best Local Similarity 27.6%; Pred. No. 0.058;
RT Matches 61; Conservative 13; Mismatches 74; Indels 73; Gaps 10;
OY 171 PRGAVPRKPVSEMPMERDRCA-----AHSLEPKKENTPDPPTSNAT-----SRAGG 218
DB 844 PPGSGGAPGPGGPGVYGERGSPGPGAGFPGARGLPPGSGGNDPPGSGSPKDG 903
OY 219 PLPP-----PSMPAVAGAAAGL-----ALLTLGVAGAGAMCWRARRA 256
DB 904 PPGAGNMGAPGSPGVGGPRDAGQPGPKSPGAGPRGAPPLGLGIGIGARGLAPPG 963
OY 257 KPSSRRHPG-----GSFRGGSLGLGG-GGAKGP-----REAPPEGLIAL 297
DB 964 MPGRGSGPGGVGVGSGKPGANGLSGERPPPGQGLPGLAGTAGEDPDGNDGSDGLPG 1023
OY 298 RGG-----GAADPPFCFHYEKVSGDYGHFVTVDDGP 330
DB 1024 RDGSPGKGRGNGSPGAP-----GAPGHP-----GFP 1052
RESULT 43
CA18_RABIT

```

ID CA18 RABIT STANDARD, PRT: 744 AA.
AC P14282;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Collagen alpha 1(VIII) chain precursor (Endothelial collagen).
GN COL8A1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89380199; PubMed=2476437;
RA Yamaguchi N., Benya P.D., van der Rest M., Ninomiya Y.;
RT "The cloning and sequencing of alpha 1(VIII) collagen cDNAs
RT demonstrate that type VIII collagen is a short chain collagen and
RT contains triple-helical and carboxyl-terminal non-triple-helical
RT domains similar to those of type X collagen."
RL J. Biol. Chem. 264:16022-16029(1989).
CC -1- FUNCTION: MAJOR COMPONENT OF THE DESCENDING MEMBRANE (BASEMENT
CC MEMBRANE) OF CORNEAL ENDOTHELIAL CELLS.
CC -1- SUBUNIT: MAY FORM HOMOTRIMERS, OR HETEROTRIMERS IN ASSOCIATION
CC WITH ALPHA 2(VIII) TYPE COLLAGENS.
CC -1- PM: PROLINS AT THE THIRD POSITION OF THE TRIPLET REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- MISCELLANEOUS: FOUR CONSECUTIVE GLY-PRO-PRO TRIPLET ARE PRESENT
CC AT THE C-TERMINUS OF THE TRIPLE-HELICAL REGION. THESE MAY PROVIDE
CC HIGH THERMAL STABILITY OF THIS REGION OF THIS REGION.
CC -1- SIMILARITY: STRONG, TO ALPHA 2 TYPES VIII AND X COLLAGENS.
CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J05042; AAA31204.1; -.
DR PIR: A34246; A34246.
DR InterPro: IPR001073; C1Q.
DR InterPro: IPR000087; Collagen.
DR Pfam: PF00386; C1Q; 1.
DR Pfam: PF00391; Collagen; 8.
DR PRINTS: PRO0007; COMPLEMENTC1Q.
DR SMART: SW00110; C1Q; 1.
DR PROSITE: PS01113; C1Q; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Cell adhesion; Collagen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 744 COLLAGEN ALPHA 1(VIII) CHAIN.
FT DOMAIN 26 117 NONHELICAL REGION (NC2).
FT DOMAIN 118 571 TRIPLE-HELICAL REGION.
FT DOMAIN 572 744 NONHELICAL REGION (NC1).
FT DOMAIN 609 744 C1Q.
SQ SEQUENCE 744 AA; 73358 MW; 2A8CFE8274E99 CRC64;

Query Match 7.6%; Score 141; DB 1; Length 744;
Best Local Similarity 23.6%; Pred. No. 0.032;
Matches 83; Conservative 29; Mismatches 108; Indels 132; Gaps 19;

```

```

DB 273 KGVTFPGPQPLKPPPGPQGPGLGVPGVGSPGLPGVGPQGDGPGQGPFG 332
QY 211 ATSGAGBPLPPSPMPAVA-----GAAGGALLLLGVAGAGAMCRRRR-----A 256
DB 333 KBOQGLPGLRGPGLPGVKGPGFPBKPD-----RGIGVPAALPRGKGKGVGAPGKG 387
QY 257 KPESRHPG-PGSPRGSGSLG-----GGGGMGPR-----EAPDELGIARGG--GAAD 304
DB 388 PGEPLPGLPMPWPFGAIGFPKKGSGGIVPGGPFGKGPGLQGPFGKPGFLGEVG 447
QY 305 PP---FPPHYKVSDDYGH-----PVYIVQGGP-----QSPPT 336
DB 448 PGIINGLPPIIPKPKBAHKGGLPGLPGVPGLLGPRGPGIIPDGLQGPFGI 499

RESULT 44
ID CA18 MOUSE STANDARD, PRT: 743 AA.
AC 000780; 09D2V4;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Collagen alpha 1(VIII) chain precursor.
GN COL8A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92362626; PubMed=1499564;
RA Murgaki Y., Shita C., Inoue M., Ooshima A., Olsen B.R.;
RA Ninomiya Y.;
RA "Alpha 1(VIII)-collagen gene transcripts encode a short-chain
RA collagen polypeptide and are expressed by various epithelial,
RA endothelial and mesenchymal cells in newborn mouse tissues."
RL Eur. J. Biochem. 207:895-902(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Iehi Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleschmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo T., Nikaido I., Pesole G., Quackenbush J.,
RA Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carinci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kariya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mommaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilting L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RA Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
CC -1- FUNCTION: MAJOR COMPONENT OF THE DESCENDING MEMBRANE (BASEMENT
CC MEMBRANE) OF CORNEAL ENDOTHELIAL CELLS.
CC -1- SUBUNIT: MAY FORM HOMOTRIMERS, OR HETEROTRIMERS IN ASSOCIATION
CC WITH ALPHA 2(VIII) TYPE COLLAGENS.
CC -1- TISSUE SPECIFICITY: HIGH LEVELS IN CAVIARUM, EYE & SKIN OF
CC NEWBORN MICE; ALSO IN VARIOUS EPITHELIAL, ENDOTHELIAL AND
CC MESENCHYMAL CELLS.
CC -1- PM: PROLINS AT THE THIRD POSITION OF THE TRIPLET REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- SIMILARITY: STRONG, TO ALPHA 2 TYPES VIII AND X COLLAGENS.
CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X66976; CA447387.1; -
DR EMBL: X66977; CA447387.1; JOINED.
DR EMBL: AK018742; BAB31383.1; -.
DR PIR: S23779; S23779.
DR WGI: MGI188463; Col18a1.
DR InterPro: IPR001073; C1q.
DR InterPro: IPR000087; Collagen.
DR Pfam: PF00386; C1q; 1.
DR Pfam: PF01391; Collagen; 7.
DR PRINTS: PRO0007; COMPLEMENTC1Q.
DR SMART: SM00110; C1q; 1.
DR PROSITE: PS01113; C1q; 1.
KM Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KM Glycoprotein; Cell adhesion; Collagen; Signal.
FT SIGNAL 1 28
FT CHAIN 1 743
FT DOMAIN 29 743 COLLAGEN ALPHA 1(VIII) CHAIN.
FT DOMAIN 118 571 NONHELIICAL REGION (NC2).
FT DOMAIN 572 743 TRIPLE-HELICAL REGION (COL1).
FT DOMAIN 608 743 NONHELIICAL REGION (NC1).
FT DOMAIN 6 6 C1Q.
FT CONFLICT 85 6 G-> R (IN REF. 2).
FT CONFLICT 109 85 H-> Y (IN REF. 1).
FT CONFLICT 109 109 K-> QG (IN REF. 2).
FT CONFLICT 248 248 P-> L (IN REF. 1).
FT CONFLICT 313 313 P-> A (IN REF. 2).
FT CONFLICT 324 324 IP-> SR (IN REF. 1).
FT CONFLICT 361 361 D-> H (IN REF. 1).
FT CONFLICT 361 361 T-> P (IN REF. 2).
FT CONFLICT 596 596 MPS-> NPF (IN REF. 1).
FT CONFLICT 717 719
SQ SEQUENCE 743 AA; 73454 MW; F584D85BD53897F4 CRC64;

Query Match
Best Local Similarity 21.6%; Score 140.5; DB 1; Length 743;
Matches 101; Conservative 30; Mismatches 135; Indels 209; Gaps 21;

QY 14 GALLIGLVGLVSGLSLPEYVNSANKRQAGSGYV---LYQICDRDLDCPRARRPG 69
DB 6 GPQLLLGILPILIS-----LNSV---RLIQAGAYYGIXPLPQIPQIPQIPQYQPLG 55
QY 70 ---PH-----SSPNVEFYLY-----L 83
DB 56 QQVPMPLKDGSLMKGKEMPMQYKKEYPPLQYMKETPPVPMKGVVPKKQKEVPLAS 115
QY 84 VGGAGGRCE---APPAPNLLTCDRPLDLRFTIKFOEYSNLTWGHFRSHDYIYA 139
DB 116 LRGEQGRGPRGPRGPRGPR-----PGLFGHMPG-----IX 146
QY 140 TSDGT-----REGLESLOG--GVCLITKMKLLKVGQSPRGANVRKPVSEMPNERDR 190
DB 147 GKRPQGGPGIGRPGMGKMGKMPGAK-----GEIQPKGEIGMPGIPGQ 196
QY 191 G--AAHSL---EPGKENTPGDPTSNATSGAGPLRPPSM----- 225
DB 197 GPRGPGILPRIGRPGRGLPGQPGAKG-ERGRPGRPGPRGLQPKKQKFGMPGLGLKG 255
QY 226 -PAVAGAGLALLLLGVAGAGAGCMWRRRRAXPSRRPG----- 265
DB 256 PRGHHGPRGPGVGLPGVKGPGVTGPRPGQGLIGVPGVQSPGMPGV 315
QY 266 -----PSFGGSLGAG--GGAGPRB 287
DB 316 GKRGQDGIPIQPGPRGKRGQGLPGLPGRGLBGVGRKPGPRGKGRGIGVPGVLGPR- 374
QY 288 AEPGEIGIALRGGAADPPFCPHYEKVSGDYGHVYIVDGP-----PQSPF 334

```

```

DB 375 GKXGIRGAPGCG-----PGRGPRGIRGPMGPRGALGPRPKRGKGVVGGPR 425
RESULT 45
ID CA1A HUMAN STANDARD; PRT; 680 AA.
AC 003692;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Collagen alpha 1(X) chain precursor.
GN COL10A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92109659; PubMed=1764025;
RA Thomas J.T., Cresswell C.J., Rash B., Nicolai H., Jones T.,
RA Solomon S., Grant M.E., Boot-Handford R.P.;
RT "The human collagen X gene. Complete primary translated sequence and
RT chromosomal localization."
RL Biochem. J. 280:617-623(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93012005; PubMed=1397333;
RA Reichenberger E., Beier F., Luvalle P., Olsen B.R., von der Mark K.,
RA Bertling W.M.;
RT "Genomic organization and full-length cDNA sequence of human collagen
RT X."
RL FEBS Lett. 311:305-310(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX Beier F., Lamm M.B., von der Mark K.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX Williams S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 52-680 FROM N.A.
RX MEDLINE=92267014; PubMed=1587271;
RA Ape S.S., Seldin M.F., Hayashi M., Olsen B.R.;
RT "Cloning of the human and mouse type X collagen genes and mapping of
RT the mouse type X collagen gene to chromosome 10."
RL Eur. J. Biochem. 206:217-224(1992).
RN [6]
RP SEQUENCE OF 561-666 FROM N.A.
RX MEDLINE=91243838; PubMed=2037056;
RA Ape S., Mattei M.-G., Olsen B.R.;
RT "Cloning of human alpha 1(X) collagen DNA and localization of the
RT COL10A1 gene to the q21-q22 region of human chromosome 6."
RL FEBS Lett. 282:393-396(1991).
RN [7]
RP SEQUENCE OF 547-655 FROM N.A.
RX MEDLINE=92077285; PubMed=1743401;
RA Reichenberger E., Aigner T., von der Mark K., Stoeck H., Bertling W.;
RT "In situ hybridization studies on the expression of type X collagen
RT in fetal human cartilage."
RL Dev. Biol. 148:562-572(1991).
RN [8]
RP REVIEW ON VARIANTS.
RX MEDLINE=97255959; PubMed=9101290;
RA Kivianen H., Tromp G., Prockop D.J.;
RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-
RT associated collagen (type IX), and network-forming collagen (type X)
RT cause a spectrum of diseases of bone, cartilage, and blood vessels."
RL Hum. Mutat. 9:300-315(1997).
RN [9]
RP VARIANTS SWCD ASP-598 AND PRO-614.
RX MEDLINE=94136476; PubMed=8304336;

```

RA Wallis G.A., Rash B., Sweetman W.A., Thomas J.T., Super M., Evans G.,
 RA Grant M.E., Boot-Handford R.P.;
 RT "Amino acid substitutions of conserved residues in the
 RT carboxyl-terminal domain of the alpha 1(X) chain of type X collagen
 RT occur in two unrelated families with metaphyseal chondrodysplasia
 RT type Schmid.";
 RL Am. J. Hum. Genet. 54:169-178(1994).
 RN [10]
 RP VARIANT SMCD ARG-591.
 RX MEDLINE=94272470; PubMed=8004099;
 RA McIntosh I., Abbott M.H., Warman M.L., Olsen B.R., Francomano C.A.;
 RT "Additional mutations of type X collagen confirm COL10A1 as the
 RT Schmid metaphyseal chondrodysplasia locus.";
 RL Hum. Mol. Genet. 3:303-307(1994).
 RN [11]
 RP VARIANT SMCD VAL-618.
 RX MEDLINE=95181449; PubMed=7876225;
 RA Chan D., Cole W.G., Rogers J.G., Bateman J.F.;
 RT "Type X collagen multimer assembly in vitro is prevented by a Gly618
 RT to Val mutation in the alpha 1(X) NCI domain resulting in Schmid
 RT metaphyseal chondrodysplasia.";
 RL J. Biol. Chem. 270:4558-4562(1995).
 RN [12]
 RP VARIANTS SMCD ARG-545; GLU-595; HIS-597; LYS-617; ARG-644 AND GLY-648.
 RX MEDLINE=95331767; PubMed=7607855;
 RA Bonaventure U., Chamblade F., Maroteaux P.;
 RT "Mutations in three subdomains of the carboxy-terminal region of
 RT collagen type X account for most of the Schmid metaphyseal
 RT dysplasias.";
 RL Hum. Genet. 96:58-64(1995).
 RN [13]
 RP VARIANT SMCD PRO-600.
 RX MEDLINE=96375754; PubMed=8782043;
 RA Wallis G.A., Rash B., Sykes B., Bonaventure U., Maroteaux P.,
 RA Zabel B., Wymne-Davies R., Grant M.E., Boot-Handford R.P.;
 RT "Mutations within the gene encoding the alpha 1(X) chain of type X
 RT collagen (COL10A1) cause metaphyseal chondrodysplasia type Schmid but
 RT not several other forms of metaphyseal chondrodysplasia.";
 RL J. Med. Genet. 33:450-457(1996).
 RN [14]
 RP VARIANTS SMCD GLU-18 AND ARG-18.
 RX MEDLINE=9720591; PubMed=9067753;
 RA Ikegawa S., Nakamura K., Nagano A., Haga N., Nakamura Y.;
 RT "Mutations in the N-terminal globular domain of the type X collagen
 RT gene (COL10A1) in patients with Schmid metaphyseal
 RT chondrodysplasia.";
 RL Hum. Mutat. 9:131-135(1997).
 RN [15]
 RP VARIANTS SMD GLU-595.
 RX MEDLINE=99057503; PubMed=9837818;
 RA Ikegawa S., Nishimura G., Nagai T., Hasegawa T., Ohashi H.,
 RA Nakamura Y.;
 RT "Mutation of the type X collagen gene 'COL10A1' causes
 RT spondylometaphyseal dysplasia.";
 RL Am. J. Hum. Genet. 63:1659-1662(1998).
 RN [16]
 RP VARIANT SMCD CYS-597.
 RX MEDLINE=99069781; PubMed=9852679;
 RA Sawai H., Ida A., Nakata Y., Koyama K.;
 RT "Novel missense mutation resulting in the substitution of tyrosine by
 RT cysteine at codon 597 of the type X collagen gene associated with
 RT Schmid metaphyseal chondrodysplasia.";
 RL J. Hum. Genet. 43:259-261(1998).
 CC -1- FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTROPHIC
 CC CHONDROCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE
 CC MINERALIZATION ZONES OF HYALINE CARTILAGE.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -1- DISEASE: DEFECTS IN COL10A1 ARE THE CAUSE OF SCHMID TYPE
 CC METAPHYSEAL CHONDRODYSPLASIA (SMCD), A DOMINANTLY INHERITED
 CC DISORDER OF THE OSSEOUS SKELETON. THE CARLINL. FEATURES OF THE
 CC PHENOTYPE ARE MILD SHORT STATURE, COXA VARA AND A WADDLING GAIT.

CC RADIOGRAPHY USUALLY SHOWS SCLEROSIS OF THE RIBS, FLARING OF THE
 CC METAPHYSES, AND A WIDE IRREGULAR GROWTH PLATE, ESPECIALLY OF THE
 CC KNEES.
 CC -1- DISEASE: DEFECTS IN COL10A1 ARE THE CAUSE OF SPONDYLOMETAPHYSEAL
 CC DYSPLASIA (SMD). SMD COMPRISES A HETEROGENEOUS GROUP OF HERITABLE
 CC SKELETAL DYSPLASIAS CHARACTERIZED BY MODIFICATIONS OF THE
 CC VERTEBRAL BODIES OF THE SPINE AND METAPHYSES OF THE TUBULAR BONES.
 CC -1- SIMILARITY: STRONG, TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.
 CC -1- SIMILARITY: CONTAINS 1 CIG DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X60382; CAA42933.1; -
 CC EMBL: X65120; CAA46236.1; -
 CC EMBL: X98568; CAA67178.1; -
 CC EMBL: AL121963; CAB87590.1; -
 CC EMBL: S68531; AAC60645.1; -
 CC EMBL: X58879; CAA41686.1; -
 CC EMBL: M74050; CAA61221.1; -
 CC EMBL: X72579; CAA51170.1; -
 CC EMBL: X72580; CAA51170.1; JOINED.
 CC PIR: S15826; S15826.
 CC PIR: S30086; S30086.
 CC PIR: A43901; A43901.
 CC PIR: S18249; S18249.
 CC PIR: S21856; S21856.
 CC PIR: S26396; S26396.
 CC GeneW: HGNC:2185; COL10A1.
 CC MIM: 120110; -
 CC MIM: 156500; -
 CC MIM: 184250; -
 CC InterPro: IPR001073; CIG.
 CC InterPro: IPR000877; Collagen.
 CC Pfam: PF00386; CIG; 1.
 CC Pfam: PF01391; Collagen; 8.
 CC PRINTS: PR00007; COMPLEMENTC1Q.
 CC SMART: SMO0110; C1Q; 1.
 CC PROSITE: PS01113; C1Q; 1.
 CC KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 CC Cartilage; Collagen; Signal; Disease mutation; Polymorphism.
 CC FT SIGNAL 1 18
 CC FT CHAIN 19 680
 CC FT DOMAIN 19 56
 CC FT DOMAIN 57 519
 CC FT DOMAIN 520 680
 CC FT DOMAIN 545 680
 CC FT VARIANT 18 18
 CC FT VARIANT 18 18
 CC FT VARIANT 18 18
 CC FT VARIANT 545 545
 CC FT VARIANT 591 591
 CC FT VARIANT 591 591
 CC Query Match 7.6%; Score 140; DB 1; Length 680;
 CC Best Local Similarity 29.1%; Pred. No. 0.034;
 CC Matches 57; Conservative 13; Mismatches 84; Indels 42; Gaps 9;
 CC
 CC 171 PG-----GAVPRK-----VSEWMERDRGAASLSRGKENTLG-----DPTNSA 211
 CC
 CC 70 PRHPPSPSPPKPKPGYSGPGLPSPSPSPSVKRPVGLPKGKRGKPYPKGV 129
 CC
 CC 212 TSGAGGPIPPSPMAVYAGAGLALLLLGVAGAGAMCWRRRKAPSPSRPG-PSGSG 270
 CC
 CC 130 GPAGLPGFPGPGPGIPGPAG---ISVPGKRGQGGPGAGPFPFPGKAPGVGFMNG 186

QY 271 RGGSLGIGGGGGGPREAPGELGALRGG--GAADP-----FCPHYKSGDYGH 320
 DB 187 QGKMGVYGA-PG-----RPGRGGLPGPGPGPGPGPGVGRGNGVPGCGIGRGDNGF 239
 QY 321 PYYIVDDP--PQSP 334
 DB 240 PGEMGPGRGPGPGPP 255

RESULT 46
 CAC2_HABCO STANDARD; PRT; 210 AA.
 ID CAC2_HABCO STANDARD; PRT; 210 AA.
 AC P16252;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cuticle collagen 2C (Fragment).
 GN 2C.
 OS Haemonchus contortus (Barber pole worm).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
 OC Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.
 RN NCB1_TaxID=6289;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90136718; PubMed=2615789;
 RX Shannaky L.M., Pratt D., Bolavenue R.J., Cox G.N.;
 RT "Cuticle collagen genes of Haemonchus contortus and Caenorhabditis
 elegans are highly conserved.";
 RT Mol. Biochem. Parasitol. 37:73-86(1989).
 CC -1- FUNCTION: NEMATODE CUTICLES ARE COMPOSED LARGELY OF COLLAGEN-LIKE
 CC PROTEINS. THE CUTICLE FUNCTIONS BOTH AS AN EXOSKELETON AND AS A
 CC BARRIER TO PROTECT THE WORM FROM ITS ENVIRONMENT.
 CC -1- MISCELLANEOUS: THIS PROTEIN SHOWS 4 POTENTIAL TRIPLE-HELICAL
 CC REGIONS, WHICH CONTAIN GLYCINE AS EVERY THIRD AMINO ACID.
 CC -1- MISCELLANEOUS: IN ALL NEMATODE CUTICLE COLLAGENS, THE POLYPEPTIDE
 CC CHAINS ARE COMPLEXED WITHIN THE CUTICLE BY DISULFIDE BONDS AND
 CC OTHER TYPES OF COVALENT CROSSLINKS.
 CC -1- SIMILARITY: BELONGS TO THE CUTICULAR COLLAGEN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; J04670; AAA29172.1; -;
 CC DR InterPro; IPR000087; Collagen.
 CC DR Pfam; PF01391; Collagen 3.
 CC DR Cuticle; Connective tissue; Repeat; Multigene family; Collagen.
 CC FT NON TER 1
 CC FT 1
 CC SQ SEQUENCE 210 AA; 19562 MW; E15FA9A2DF3D74B CRC64;

Query Match 7.5%; Score 139.5; DB 1; Length 210;
 Best Local Similarity 31.1%; Pred. No. 0.011; Indels 31; Gaps 5;
 Matches 50; Conservative 5; Mismatches 75;

QY 176 VPRKVSSEMPERDRGAHSLRPGKNTLPDGP--TSNATSRGASPLPSPMPAVAGAAG 233
 DB 51 IIPCKKPCQGRPPGPPGPGPGTGGNGAGANAPAPPPGPKPPGPGAGAP- 109
 QY 234 GLALLLGVAGAGAMCWRRAKPSSESHPG---PQSGRGSLGIGGGGNGPREAE 289
 DB 110 -----GAAGPAGANA-----PSEPLVPGPDPPTGTGEGGPPGNAGAPGAGAGA 155
 QY 290 PSELGIALRGGAADPPFCPHYKSGDYGPVYIVDDGP 330
 DB 156 PSEKGRGGDGHGAP-----GNAHGRCGCGGPP 186

ID CAlB MOUSE STANDARD; PRT; 1804 AA.
 AC Q6125; O64047;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Collagen alpha 1(XI) chain precursor.
 GN COL11A1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCB1_TaxID=10090;
 RX [1]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Embryo;
 RX MEDLINE=96015067; PubMed=8510046;
 RA Yoshioke H., Inoguchi K., Khaleduzzaman M., Ninomiya Y.,
 RA Andrikopoulos K., Ramirez F.;
 RT "Coding sequence and alternative splicing of the mouse alpha 1(XI)
 RT collagen gene (Col11a1).";
 RT Genomics 28:337-340(1995).
 RN [2]
 RP SEQUENCE OF 181-198 FROM N.A., AND CHONDRODYSPLASIA VARIANT.
 RC STRAIN=C57BL/6;
 RX MEDLINE=95163095; PubMed=7859283;
 RA Li Y., Lacerda D.A., Warman M.L., Belier D.R., Yoshioke H.,
 RA Ninomiya Y., Oxford J.T., Morris N.P., Andrikopoulos K.,
 RA Ramirez F., Wardell B.B., Liffert G.D., Tewscher C., Woodward S.R.,
 RA Taylor B.A., Seemuller R.E., Olsen B.R.;
 RT "A fibrillar collagen gene, Col1a1, is essential for skeletal
 RT morphogenesis.";
 RT Cell 80:423-430(1995).
 CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN FIBRILLOGENESIS BY
 CC CONTROLLING LATERAL GROWTH OF COLLAGEN II FIBRILS.
 CC -1- SUBUNIT: TRIMERS COMPOSED OF THREE DIFFERENT CHAINS: ALPHA 1(XI),
 CC ALPHA 2(XI), AND ALPHA 3(XI). ALPHA 3(XI) IS A POST-TRANSLATIONAL
 CC MODIFICATION OF ALPHA 1(XI). ALPHA 1(XI) CAN ALSO BE FOUND INSTEAD
 CC OF ALPHA 3(XI)=1(XI) (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -1- DISEASE: DEFECTS IN COL1A1 ARE ASSOCIATED WITH CHONDRODYSPLASIA,
 CC AN AUTOSOMAL RECESSIVE DISEASE CHARACTERIZED BY SKELETAL DEFECTS
 CC CAUSED BY ABNORMALITIES IN THE CARTILAGE OF LIMBS, RIBS, MANDIBLES
 CC AND TRACHEA.
 CC -1- SIMILARITY: BELONGS TO THE FIBRILLAR CLASS OF COLLAGENS.
 CC -1- SIMILARITY: HIGH TO ALPHA 1(V) AND ALPHA 3(V) CHAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; D38162; BAA07367.1; -;
 CC DR EMBL; S74574; AAB33439.1; -;
 CC DR MGI; 88446; Col1a1.
 CC DR InterPro; IPR000087; Collagen.
 CC DR InterPro; IPR000085; Fib collagen_C.
 CC DR InterPro; IPR001791; LamInG.
 CC DR InterPro; IPR001230; Prenyl_site.
 CC DR InterPro; IPR003129; TSPN.
 CC DR Pfam; PF01391; Collagen; 16.
 CC DR Pfam; PF01410; COLF1; 1.
 CC DR Pfam; PF02210; TSPN; 1.
 CC DR ProDom; PD000007; Collagen; 1.
 CC DR ProDom; PD002078; Fib collagen_C; 1.
 CC DR SMART; SM00038; COLF1; 1.
 CC DR SMART; SM00282; LamG; 1.
 CC DR SMART; SM00210; TSPN; 1.
 CC KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;

ID CA2B_HUMAN STANDARD; PRT; 1736 AA.
 AC P13942; Q13273; Q13271; Q13272; Q07751; Q93866; Q9UIP9;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Collagen alpha 2(XI) chain precursor.
 GN COL11A2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96032717; PubMed=7559422;
 RA Vuorio M.M., Pihlajama T., Vandenbergh P., Prockop D.J.,
 RA Ala-Kokko L.;
 RT "The human COL11A2 gene structure indicates that the gene has not
 RT evolved with the genes for the major fibrillar collagens.";
 RL J. Biol. Chem. 270:22873-22881(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Tubby B.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 59-807 FROM N.A.
 RC TISSUE=Cartilage;
 RX MEDLINE=93314795; PubMed=8325374;
 RA Zhidkova N.I., Brewton R.G., Wayne R.;
 RA "Molecular cloning of PAPP (proline/arginine-rich protein) from human
 RT cartilage and subsequent demonstration that PAPP is a fragment of the
 RT NH2-terminal domain of the collagen alpha 2(XI) chain.";
 RL Peps Lett. 326:25-26(1993).
 RN [4]
 RP SEQUENCE OF 730-1690 FROM N.A.
 RX MEDLINE=89340485; PubMed=2760050;
 RA Kimura T., Cheah K.S.E., Chan S.D.H., Lui V.C.H., Mattei M.-G.,
 RA van der Reet M., Ono K., Solomon E., Ninomiya Y., Olsen B.R.;
 RT "The human alpha 2(XI) collagen (COL11A2) chain. Molecular cloning of
 RT cDNA and genomic DNA reveals characteristics of a fibrillar collagen
 RT with differences in genomic organization.";
 RL J. Biol. Chem. 264:13910-13916(1989).
 RN [5]
 RP SEQUENCE OF 1-537 FROM N.A.
 RX MEDLINE=96435918; PubMed=8838804;
 RA Lui V.C., Ng L.J., Sat E.W., Cheah K.S.;
 RT "The human alpha 2(XI) collagen gene (COL11A2): completion of coding
 RT information, identification of the promoter sequence, and precise
 RT localization within the major histocompatibility complex reveal
 RT overlap with the KES gene.";
 RL Genomics 32:401-412(1996).
 RN [6]
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=95238468; PubMed=7721876;
 RA Zhidkova N.I., Justice S.K., Wayne R.;
 RT "Alternative mRNA processing occurs in the variable region of the
 RT pro-alpha 1(XI) and pro-alpha 2(XI) collagen chains.";
 RL J. Biol. Chem. 270:9486-9493(1995).
 RN [7]
 RP DISEASE.
 RX PubMed=10677296;
 RA Melkonian M., Brunner H.G., Manouvrier S., Hennekam R.,
 RA Superti-Furga A., Kaeaeilaenen H., Pauli R.M., van Essen T.,
 RA Warman M.L., Bonaventure J., Mlyn P., Ala-Kokko L.;
 RT "Autosomal recessive osteochondrodysplasia is
 RT associated with loss-of-function mutations in the COL11A2 gene.";
 RL Am. J. Hum. Genet. 66:368-377(2000).
 RN [8]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=97253959; PubMed=9101290;
 RA Kulvanenti H., Tromp G., Prockop D.J.;
 RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-
 RT associated collagen (type IX), and network-forming collagen (type X)
 RT cause a spectrum of diseases of bone, cartilage, and blood vessels.";

RL Hum. Mutat. 9:300-315(1997).
 RN [9]
 RP VARIANT OSMED ARG-661.
 RX MEDLINE=95163096; PubMed=7859284;
 RA Vlkula M., Mariman E.C.M., Lui V.C.H., Zhidkova N.I., Tiller G.E.,
 RA Goldring M.B., van Beersum S.B.C., de Waal Malefijt M.C.,
 RA van den Hoogen F.H.J., Roberts H.-H., Wayne R., Cheah K.S.E.,
 RA Olsen B.R., Warman M.L., Brunner H.G.;
 RT "Autosomal dominant and recessive osteochondrodysplasias associated
 RT with the COL11A2 locus.";
 RL Cell 80:431-437(1995).
 RN [10]
 RP VARIANTS GLY-593; LYS-824; LEU-879; THR-1316 AND GLN-1600.
 RX PubMed=9585596;
 RA Koga H., Sakou T., Taketomi E., Hayashi K., Numasawa T., Harata S.,
 RA Yone K., Matsunaga S., Ohterud B., Inoue I., Leppert M.;
 RT "Genetic mapping of ossification of the posterior longitudinal
 RT ligament of the spine.";
 RL Am. J. Hum. Genet. 62:1460-1467(1998).
 RN [11]
 RP VARIANT W2S GLU-955.
 RX PubMed=9805126;
 RA Pihlajama T., Prockop D.J., Faber J., Winterpacht A., Zabel B.,
 RA Gledhill A., Wiesbauer P., Spranger J., Ala-Kokko L.;
 RT "Heterozygous glycine substitution in the COL11A2 gene in the original
 RT patient with the Weissenbacher-Zwemmelier syndrome demonstrates its
 RT identity with heterozygous OSMD (homocystin-like syndrome).";
 RL Am. J. Med. Genet. 80:115-120(1998).
 RN [12]
 RP VARIANT STU3 940-GLY--PRO-948 DEL.
 RX PubMed=9506662;
 RA Strko-Osadek D.A., Murray M.A., Scott J.A., Lavery M.A., Warman M.L.,
 RA Rodin N.H.;
 RT "Stickler syndrome without eye involvement is caused by mutations in
 RT COL11A2, the gene encoding the alpha-2(XI) chain of type XI
 RT collagen.";
 RL J. Pediatr. 132:368-371(1998).
 RN [13]
 RP VARIANTS DPM13 GLU-808 AND CYS-1034, AND REVISIONS TO 1031-1032.
 RX MEDLINE=20047768; PubMed=10581026;
 RA McGulir W.T., Praead S.D., Griffith A.J., Kunst H.P.M., Green G.E.,
 RA Shargel K.B., Runge C., Huybrechts C., Mueller R.F., Lynch E.,
 RA King M.-C., Brunner H.G., Creemers C.W.R.J., Takemura M., Li S.-W.,
 RA Arta M., Wayne R., Prockop D.J., Van Camp G., Smith R.J.H.;
 RT "Mutations in COL11A2 cause non-syndromic hearing loss (DPM13).";
 RL Nat. Genet. 23:413-419(1999).
 RN [14]
 RP FUNCTION: MAY PLAY AN IMPORTANT ROLE IN FIBRINOGENESIS BY
 CC CONTROLLING LATERAL GROWTH OF COLLAGEN II FIBRILS.
 CC -1 SUBUNIT: TRIMERS COMPOSED OF THREE DIFFERENT CHAINS: ALPHA 1(XI),
 CC ALPHA 2(XI), AND ALPHA 3(XI). ALPHA 3(XI) IS A POST-TRANSLATIONAL
 CC MODIFICATION OF ALPHA 1(XI). ALPHA 3(XI) CAN ALSO BE FOUND INSTEAD
 CC OF ALPHA 3(XI)=1(III).
 CC -1 ALTERNATIVE PRODUCTS: 8 isoforms; 1 (shown here), 2, 3, 4, 5, 6, 7
 CC and 8; may be produced by alternative splicing. They lack exons 6,
 CC 7 or 8 or a combination of these exons.
 CC -1 PPM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -1 PPM: A DISULFIDE-BONDED PEPTIDE CALLED PROLINE/ARGININE-RICH
 CC PROTEIN OR PAPP IS RELEASED FROM THE AMINO TERMINUS DURING
 CC EXTRACELLULAR PROCESSING AND IS SUBSEQUENTLY RETAINED IN THE
 CC CARTILAGE MATRIX FROM WHICH IT CAN BE ISOLATED IN SIGNIFICANT
 CC AMOUNTS.
 CC -1 DISEASE: Defects in COL11A2 are the cause of Stickler syndrome
 CC type 3 (ST3). It is an autosomal dominant disorder characterized
 CC by oro-facial, auditory and skeletal manifestations, such as
 CC midfacial hypoplasia, cleft palate, osteoarthritis, and
 CC sensorineural hearing loss. Differently from Stickler syndrome
 CC type 1 and 2, no ocular involvement is observed. This disorder is
 CC also referred to as Stickler-like syndrome or non-ocular Stickler
 CC syndrome.
 CC -1 DISEASE: Defects in COL11A2 are the cause of autosomal recessive
 CC osteochondrodysplasia (OSMED), a skeletal dysplasia
 CC accompanied by severe hearing loss. The phenotype overlaps that of

```

CC autosomal dominant skeletal disorders (Stickler and Marshall
CC syndromes) but can be distinguished by disproportionately short
CC limbs and lack of ocular involvement.
CC -1- DISEASE: Defects in COL1A2 are the cause of Weissenbacher-
CC Zeymüller syndrome (WZS), an autosomal dominant disorder allelic
CC with STL3 and OSMD. It is also referred to as heterozygous OSMD.
CC -1- DISEASE: Defects in COL1A2 are the cause of autosomal dominant
CC nonsyndromic sensorineural deafness type 13 (DFNA13). Affected
CC individuals experience progressive hearing loss beginning in the
CC second to fourth decades, eventually making use of amplification
CC mandatory.
CC -1- SIMILARITY: BELONGS TO THE FIBRILLAR CLASS OF COLLAGENS.
CC -1- DATABASE: NAME=Hereditary hearing loss homepage;
CC NOTE=Gene page;
CC WWW="http://www.uia.ac.be/dnalab/hhb/hhgenes.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32169; AAC50214.1; -
CC EMBL; U32169; AAC50213.1; -
CC EMBL; U32169; AAC50215.1; -
CC EMBL; AL031228; CAA30240.1; -
CC EMBL; L19887; AAA35498.1; -
CC EMBL; U04874; AAA52034.1; -
CC EMBL; U41069; AAC17464.1; -
CC EMBL; U41065; AAC17464.1; JOINED.
CC EMBL; U41066; AAC17464.1; JOINED.
CC EMBL; U41067; AAC17464.1; JOINED.
CC PIR; A32645; A32645.
CC Genew; HGNC:2187; COL1A2.
CC MIM; 120290; -
CC MIM; 184840; -
CC MIM; 215150; -
CC MIM; 277610; -
CC MIM; 601868; -
CC InterPro; IPR000087; Collagen.
CC InterPro; IPR000885; Fib_collagen_C.
CC InterPro; IPR001791; Laminin_G.
CC InterPro; IPR001230; Prenyl_site.
CC InterPro; IPR003129; TSPN.
CC Pfam; PF01391; Collagen; 18.
CC Pfam; PF01410; COLFR; 1.
CC Pfam; PF02210; TSPN; 1.
CC ProDom; PD000007; Collagen; 1.
CC ProDom; PD002078; Fib_collagen_C; 1.
CC SMART; SMO0038; COLFR; 1.
CC SMART; SMO0282; LamG; 1.
CC SMART; SMO0210; TSPN; 1.
Query Match 7.5%; Score 138.5; DB 1; Length 1736;
Best Local Similarity 31.7%; Pred. No. 0.11; Indels 49; Gaps 14;
Matches 64; Conservative 18; Mismatches 71;

```

```

DB 1482 KBEKVGQSPGHP-----GPP 1497
RESULT 50
CA24 CAEEL STANDARD; PRT; 1758 AA.
ID CA24 CAEEL
PI 17140; Q19098; Q19099;
DT 01-AUG-1990 (Rel. 15; Created)
DT 01-OCT-1994 (Rel. 30; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Collagen alpha 2(IV) chain precursor (lethal protein 2).
GN LET-2 OR CLB-1 OR F01G12.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoa; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCB1_TaxID=6239;
OX NCB1
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RC STRAIN=Bristol N2;
RX MEDLINE=94012964; PubMed=7691828;
RA Sibley M.H., Johnson J.J., Mello C.C., Kramer J.M.;
RT "Genetic identification, sequence, and alternative splicing of the
RT Caenorhabditis elegans alpha 2(IV) collagen gene."
RL J. Cell Biol. 123:255-264(1993).
RN [2]
RP PRELIMINARY SEQUENCE OF 1495-1758 FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=90008929; PubMed=2793871;
RA Guo X., Kramer J.M.;
RT "The two Caenorhabditis elegans basement membrane (type IV) collagen
RT genes are located on separate chromosomes."
RL J. Biol. Chem. 264:17574-17582(1989).
RN [3]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC STRAIN=Bristol N2;
RA Wu X., Le T.T.;
RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
RN [4]
RP VARIANTS.
RX MEDLINE=94320591; PubMed=8045258;
RA Sibley M.H., Graham P.L., von Mendel N., Kramer J.M.;
RT "Mutations in the alpha 2(IV) basement membrane collagen gene of
RT Caenorhabditis elegans produce phenotypes of differing severities."
RL EMO J. 13:3278-3285(1994).
CC -1- FUNCTION: Collagen type IV is specific for basement membranes.
CC Vital for embryonic development.
CC -1- SUBUNIT: TRIMERS OF TWO ALPHA 1(IV) AND ONE ALPHA 2(IV) CHAIN.
CC TYPE IV COLLAGEN FORMS A MESH-LIKE NETWORK LINKED THROUGH
CC INTERMOLECULAR INTERACTIONS BETWEEN 7S DOMAINS AND BETWEEN NCI
CC DOMAINS.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms, I/a (shown here) and II/b; are
CC produced by alternative splicing.
CC -1- DEVELOPMENTAL STAGE: Isoform I is predominant in embryos and
CC isoform II is predominant in the larvae and adults.
CC -1- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
CC DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
CC G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
CC CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
CC TRIPLE-HELICAL 7S DOMAIN.
CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPLEPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH
CC ARE INVOLVED IN INTRA- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF
CC THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE
CC IV COLLAGENS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

